```
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (653)
;   OTHER INFORMATION: Unidentified at time of filing
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (656)
;   OTHER INFORMATION: Unidentified at time of filing
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (660)
;   OTHER INFORMATION: Unidentified at time of filing
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (661)
;   OTHER INFORMATION: Unidentified at time of filing
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (662)
;   OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-4

  Query Match             93.1%;  Score 27;  DB 3;  Length 888;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              |||||:
Db        248 SSGPSI 253


RESULT 13
US-08-658-857B-16
; Sequence 16, Application US/08658857B
; Patent No. 6040435
;  GENERAL INFORMATION:
;    APPLICANT:  Hancock, Robert E. W.
;    APPLICANT:  Karunaratne, Nedra
;    TITLE OF INVENTION:  ANTIMICROBIAL CATIONIC PEPTIDES
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,857B
;      FILING DATE:  May 31, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/460,464
;      FILING DATE:  June 2, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07420/014001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-658-857B-16

  Query Match             89.7%;  Score 26;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 46;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||:|
Db         12 SSGPAL 17


RESULT 14
US-08-763-226C-16
; Sequence 16, Application US/08763226C
; Patent No. 6057291
;  GENERAL INFORMATION:
;    APPLICANT:  Hancock, Robert E. W.
;    APPLICANT:  Karunaratne, Nedra
;    TITLE OF INVENTION:  ANTIMICROBIAL CATIONIC PEPTIDES
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/763,226C
;      FILING DATE:  10-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/658,857
;      FILING DATE:  31-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07420/014001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-763-226C-16

  Query Match             89.7%;  Score 26;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 46;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||:|
Db         12 SSGPAL 17


RESULT 15
US-09-307-200-16
; Sequence 16, Application US/09307200
; Patent No. 6297215
;  GENERAL INFORMATION:
;    APPLICANT:  Hancock, Robert E. W.
;    APPLICANT:  Karunaratne, Nedra
;    TITLE OF INVENTION:  ANTIMICROBIAL CATIONIC PEPTIDES
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
```

```
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows95
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/307,200
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/763,226
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Haile, Lisa A.
;       REGISTRATION NUMBER:  38,347
;       REFERENCE/DOCKET NUMBER:  07420/014001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619/678-5070
;       TELEFAX:  619/678-5099
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-307-200-16

  Query Match             89.7%;  Score 26;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 46;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||:|
Db         12 SSGPAL 17


Search completed: March 10, 2004, 09:28:53
Job time : 3.32296 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 26.7237 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	185	99.5	34	9	US-09-848-834A-10	Sequence 10, Appl
2	185	99.5	50	9	US-09-848-834A-18	Sequence 18, Appl
3	119	64.0	194	14	US-10-295-074-46	Sequence 46, Appl
4	118	63.4	158	14	US-10-297-942-12	Sequence 12, Appl
5	116	62.4	285	14	US-10-295-074-11	Sequence 11, Appl
6	116	62.4	287	14	US-10-295-074-13	Sequence 13, Appl
7	114	61.3	158	14	US-10-297-942-10	Sequence 10, Appl
8	112	60.2	21	9	US-09-943-548-3	Sequence 3, Appli
9	112	60.2	21	9	US-09-848-834A-4	Sequence 4, Appli
10	112	60.2	21	9	US-09-785-215-6	Sequence 6, Appli
11	112	60.2	21	10	US-09-405-986-2	Sequence 2, Appli
12	112	60.2	21	14	US-10-204-362-6	Sequence 6, Appli
13	112	60.2	21	14	US-10-339-522-3	Sequence 3, Appli
14	112	60.2	21	14	US-10-223-711-8	Sequence 8, Appli
15	112	60.2	21	14	US-10-223-809A-6	Sequence 6, Appli
16	112	60.2	21	14	US-10-261-208-5	Sequence 5, Appli
17	112	60.2	21	14	US-10-295-074-5	Sequence 5, Appli
18	112	60.2	21	15	US-10-372-111-8	Sequence 8, Appli
19	112	60.2	32	15	US-10-411-544-28	Sequence 28, Appl
20	112	60.2	37	9	US-09-848-834A-14	Sequence 14, Appl
21	112	60.2	158	14	US-10-297-942-2	Sequence 2, Appli
22	112	60.2	158	14	US-10-297-942-18	Sequence 18, Appl
23	112	60.2	158	14	US-10-297-942-20	Sequence 20, Appl
24	112	60.2	194	14	US-10-295-074-47	Sequence 47, Appl
25	112	60.2	285	14	US-10-295-074-9	Sequence 9, Appli
26	112	60.2	287	14	US-10-295-074-15	Sequence 15, Appl
27	112	60.2	441	15	US-10-452-024-180	Sequence 180, App
28	112	60.2	441	15	US-10-452-024-183	Sequence 183, App
29	112	60.2	444	15	US-10-452-024-185	Sequence 185, App
30	112	60.2	451	15	US-10-452-024-186	Sequence 186, App
31	112	60.2	452	15	US-10-452-024-184	Sequence 184, App
32	112	60.2	463	10	US-09-816-467-2	Sequence 2, Appli
33	112	60.2	469	15	US-10-452-024-182	Sequence 182, App
34	112	60.2	472	15	US-10-452-024-181	Sequence 181, App
35	112	60.2	514	14	US-10-295-074-49	Sequence 49, Appl
36	112	60.2	514	14	US-10-295-074-51	Sequence 51, Appl
37	112	60.2	514	14	US-10-295-074-59	Sequence 59, Appl
38	112	60.2	517	14	US-10-295-074-53	Sequence 53, Appl
39	112	60.2	573	15	US-10-452-024-177	Sequence 177, App
40	112	60.2	605	14	US-10-130-973A-11	Sequence 11, Appl
41	112	60.2	665	14	US-10-130-973A-9	Sequence 9, Appli
42	112	60.2	882	14	US-10-130-973A-3	Sequence 3, Appli
43	112	60.2	907	14	US-10-130-973A-5	Sequence 5, Appli
44	112	60.2	1052	14	US-10-130-973A-17	Sequence 17, Appl
45	112	60.2	1112	14	US-10-130-973A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-10
; Sequence 10, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of t
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylysin) linked by a spacer to a
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated phenylalanine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylysin)
; NAME/KEY: PEPTIDE
; LOCATION: (22)..(25)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (26)..(34)
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (34)..(34)
; OTHER INFORMATION: Amidated glycine or glycinamide

```
US-09-848-834A-10

  Query Match             99.5%;  Score 185;  DB 9;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-18;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||||||||||||||
Db          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33


RESULT 2
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer t
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match             99.5%;  Score 185;  DB 9;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-17;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||||||||||||||
Db         17 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49


RESULT 3
US-10-295-074-46
; Sequence 46, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 194
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (110)..(124)
;   OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (125)..(145)
;   OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (2)..(109)
;   OTHER INFORMATION: hTNF amino acids 1-108
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (146)..(194)
;   OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

  Query Match             64.0%;  Score 119;  DB 14;  Length 194;
  Best Local Similarity   81.5%;  Pred. No. 3.6e-08;
  Matches   22;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHW 27
              |||||||||||||||||||||  : |
Db        125 FNNFTVSFWLRVPKVSASHLEAEAKPW 151


RESULT 4
US-10-297-942-12
; Sequence 12, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-12

  Query Match             63.4%;  Score 118;  DB 14;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 4e-08;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEG 22
              ||||||||||||||||||||||
Db        133 FNNFTVSFWLRVPKVSASHLEG 154


RESULT 5
US-10-295-074-11
; Sequence 11, Application US/10295074
```

```
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 285
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: 2 human IL5 monomers joined by P2 and P30 epitopes
US-10-295-074-11

  Query Match             62.4%;  Score 116;  DB 14;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. 1.4e-07;
  Matches   22;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPS 24
              |||||||||||||||||||||| |:
Db        150 FNNFTVSFWLRVPKVSASHLEIPT 173


RESULT 6
US-10-295-074-13
; Sequence 13, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 287
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Two human IL5 monomers joined by diglycine linker and including t
;   OTHER INFORMATION: erminally positioned P30 and P2 epitopes
US-10-295-074-13

  Query Match             62.4%;  Score 116;  DB 14;  Length 287;
  Best Local Similarity   91.7%;  Pred. No. 1.4e-07;
  Matches   22;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPS 24
              |||||||||||||||||||||| |:
Db         24 FNNFTVSFWLRVPKVSASHLEIPT 47


RESULT 7
US-10-297-942-10
; Sequence 10, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-10

  Query Match             61.3%;  Score 114;  DB 14;  Length 158;
  Best Local Similarity   70.6%;  Pred. No. 1.4e-07;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    8;  Gaps    1;

Qy          1 FNNFTVSFWLRVPKVSASHLEG--------PSLH 26
              ||||||||||||||||||||||          || |
Db         41 FNNFTVSFWLRVPKVSASHLEQVLFKGQGCPSTH 74


RESULT 8
US-09-943-548-3
; Sequence 3, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
;  APPLICANT: Rittershaus, Charles W.
;  APPLICANT:  Thomas, Lawrence J.
;  TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVIT
;  FILE REFERENCE: TCS-411.1P US-1; Tcs-411.1P US-2
;  CURRENT APPLICATION NUMBER: US/09/943,548
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 08/432,483
;  PRIOR FILING DATE: 1995-05-01
;  PRIOR APPLICATION NUMBER: PCT/US96/06147
;  PRIOR FILING DATE: 1996-05-01
;  PRIOR APPLICATION NUMBER: 08/945,289
;  PRIOR FILING DATE: 1997-10-17
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-09-943-548-3

  Query Match             60.2%;  Score 112;  DB 9;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3e-08;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |||||||||||||||||||||
Db          1 FNNFTVSFWLRVPKVSASHLE 21


RESULT 9
US-09-848-834A-4
; Sequence 4, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Tetanus bacillus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(21)
;   OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus
;   OTHER INFORMATION: Toxoid Precursor (Tentoxylysin)
```

US-09-848-834A-4

Query Match 60.2%; Score 112; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10
US-09-785-215-6
; Sequence 6, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-785-215-6

Query Match 60.2%; Score 112; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11
US-09-405-986-2
; Sequence 2, Application US/09405986
; Publication No. US20030157115A1
; GENERAL INFORMATION:
; APPLICANT: BAY, Sylvie
; APPLICANT: CANTACUZENE, Daniele
; APPLICANT: LECLERC, Claude
; APPLICANT: LO-MAN, Richard
; TITLE OF INVENTION: MULTIPLE ANTIGEN GLYCOPEPTIDE CARBOHYDRATE, VACCINE
; TITLE OF INVENTION: COMPRISING THE SAME AND USE THEREOF
; FILE REFERENCE: 134I US 3565
; CURRENT APPLICATION NUMBER: US/09/405,986
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: 60/041,726
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-405-986-2

Query Match 60.2%; Score 112; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12
US-10-204-362-6
; Sequence 6, Application US/10204362
; Publication No. US20030086938A1
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; TITLE OF INVENTION: No. US20030086938A1el Method For Down-Regulation Of Amyloid
; FILE REFERENCE: 3631-0120P
; CURRENT APPLICATION NUMBER: US/10/204,362
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-204-362-6

Query Match 60.2%; Score 112; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
US-10-339-522-3
; Sequence 3, Application US/10339522
; Publication No. US20030108559A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-3
; CURRENT APPLICATION NUMBER: US/10/339,522
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-10-339-522-3

Query Match 60.2%; Score 112; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 14
US-10-223-711-8
; Sequence 8, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Pravin T.P.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

```
;  FILE REFERENCE: 18525/04058
;  CURRENT APPLICATION NUMBER: US/10/223,711
;  CURRENT FILING DATE:  2002-08-19
;  PRIOR APPLICATION NUMBER: 09/148,711
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 08/460,502
;  PRIOR FILING DATE: 1995-06-02
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
US-10-223-711-8

  Query Match             60.2%;  Score 112;  DB 14;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3e-08;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |||||||||||||||||||||
Db          1 FNNFTVSFWLRVPKVSASHLE 21


RESULT 15
US-10-223-809A-6
; Sequence 6, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  APPLICANT:  Rasmussen, Peter Birk et al.
;  TITLE OF INVENTION: No. US20030157117A1el Method for Down-Regulation of Amyloid
;  FILE REFERENCE: 674542-2008
;  CURRENT APPLICATION NUMBER: US/10/223,809A
;  CURRENT FILING DATE:  2002-08-20
;  PRIOR APPLICATION NUMBER: US 60/337,543
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: US 60/373,027
;  PRIOR FILING DATE: 2002-04-16
;  PRIOR APPLICATION NUMBER: DE 2001 01231
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: DE 2002 0058
;  PRIOR FILING DATE: 2002-04-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
US-10-223-809A-6

  Query Match             60.2%;  Score 112;  DB 14;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3e-08;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |||||||||||||||||||||
Db          1 FNNFTVSFWLRVPKVSASHLE 21


Search completed: March 10, 2004, 10:25:48
Job time : 26.7237 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 33.6031 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	112	60.2	451	2	Q9LA13	Q9la13 clostridium
2	112	60.2	1310	2	Q93N27	Q93n27 clostridium
3	62	33.3	1268	2	Q45851	Q45851 clostridium
4	61	32.8	1278	2	Q57236	Q57236 clostridium
5	58	31.2	361	2	Q45848	Q45848 clostridium
6	58	31.2	361	2	Q45846	Q45846 clostridium
7	58	31.2	441	2	Q9X708	Q9x708 clostridium
8	58	31.2	1291	2	Q9ZAJ8	Q9zaj8 clostridium
9	58	31.2	1291	2	Q93G71	Q93g71 clostridium
10	58	31.2	1291	2	Q933K0	Q933k0 clostridium
11	58	31.2	1291	2	Q08077	Q08077 clostridium
12	58	31.2	1291	2	Q8GR96	Q8gr96 clostridium
13	57	30.6	430	2	Q9XAV1	Q9xav1 pseudomonas
14	57	30.6	502	16	Q9X8T8	Q9x8t8 streptomyce
15	56.5	30.4	367	2	Q45861	Q45861 clostridium
16	56.5	30.4	367	2	Q45862	Q45862 clostridium
17	56.5	30.4	1252	2	Q8KZM3	Q8kzm3 clostridium
18	56.5	30.4	1255	2	Q9FAR6	Q9far6 clostridium
19	56	30.1	804	10	Q8L853	Q81853 arabidopsis
20	56	30.1	812	10	O64620	O64620 arabidopsis
21	56	30.1	1251	2	Q9K395	Q9k395 clostridium
22	55	29.6	1280	2	Q9ZAJ5	Q9zaj5 clostridium
23	54.5	29.3	476	16	Q8PM68	Q8pm68 xanthomonas
24	54.5	29.3	1285	2	Q45967	Q45967 clostridium
25	54.5	29.3	1285	2	Q9LBR1	Q9lbr1 clostridium
26	54.5	29.3	1291	2	Q93HT3	Q93ht3 clostridium
27	54	29.0	91	13	Q9PRH0	Q9prh0 anguilla ja
28	53	28.5	606	16	Q8XIK6	Q8xik6 clostridium
29	52	28.0	51	16	Q88GE4	Q88ge4 pseudomonas
30	52	28.0	202	16	Q829M1	Q829m1 streptomyce
31	52	28.0	466	12	Q8JXJ9	Q8jxj9 tomato spot
32	52	28.0	466	12	Q8JXK0	Q8jxk0 tomato spot
33	52	28.0	467	12	Q8JXK2	Q8jxk2 tomato spot
34	52	28.0	467	12	O37369	O37369 tomato spot
35	52	28.0	467	12	O37367	O37367 tomato spot
36	52	28.0	467	12	Q8JVL0	Q8jvl0 tomato spot
37	52	28.0	467	12	Q8JXK4	Q8jxk4 tomato spot
38	51.5	27.7	317	4	Q9Y2R8	Q9y2r8 homo sapien
39	51.5	27.7	345	16	Q8VRF4	Q8vrf4 bradyrhizob
40	51.5	27.7	381	2	Q8VKZ1	Q8vkz1 enterobacte
41	51.5	27.7	503	8	Q7YKK8	Q7ykk8 utricularia
42	51.5	27.7	503	8	Q7YKK7	Q7ykk7 utricularia
43	51.5	27.7	520	4	Q8WVD0	Q8wvd0 homo sapien
44	51.5	27.7	1127	16	Q8G815	Q8g815 bifidobacte
45	51	27.4	130	2	Q93PY5	Q93py5 pseudomonas

ALIGNMENTS

RESULT 1
Q9LA13
ID Q9LA13 PRELIMINARY; PRT; 451 AA.
AC Q9LA13;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF154828; AAF73267.1; -.
DR HSSP; P04958; 1A8D.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz_legume.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT NON_TER 1 1
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 60.2%; Score 112; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||
Db 83 FNNFTVSFWLRVPKVSASHLE 103

RESULT 2
Q93N27

```
ID   Q93N27      PRELIMINARY;      PRT;  1310 AA.
AC   Q93N27;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tetanus toxin (Fragment).
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shumin Z., Dianliang L.;
RT   "Cloning and sequence analysis of tetanus toxin gene.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389424; AAK72964.2; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
FT   NON_TER       1      1
FT   NON_TER    1310   1310
SQ   SEQUENCE   1310 AA;  150316 MW;  9EADDC914418E450 CRC64;

  Query Match             60.2%;  Score 112;  DB 2;  Length 1310;
  Best Local Similarity   100.0%;  Pred. No. 5e-07;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |||||||||||||||||||||
Db        948 FNNFTVSFWLRVPKVSASHLE 968


RESULT 3
Q45851
ID   Q45851      PRELIMINARY;      PRT;  1268 AA.
AC   Q45851;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin type F.
GN   BONT /F.
OS   Clostridium baratii.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1561;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93252228; PubMed=8486245;
RA   Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA   Richardson P.T.;
RT   "Nucleotide sequence of the gene coding for Clostridium barati type F
RT   neurotoxin: Comparison with other clostridial neurotoxins.";
RL   FEMS Microbiol. Lett. 108:175-182(1993).
DR   EMBL; X68262; CAA48329.1; -.
DR   PIR; S33411; S33411.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1268 AA;  145513 MW;  963040091AC15ED2 CRC64;

  Query Match             33.3%;  Score 62;  DB 2;  Length 1268;
  Best Local Similarity   64.3%;  Pred. No. 6.5;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              : ||:||||:|:||
Db        922 YQNFSVSFWVRIPK 935


RESULT 4
Q57236
ID   Q57236      PRELIMINARY;      PRT;  1278 AA.
AC   Q57236; Q45863;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type F (BONT/F protein).
GN   BONT/F.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 10281;
RA   Hutson R.A., Collins M.D.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 635-1000 FROM N.A.
RC   STRAIN=NCTC 1028;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-27 FROM N.A.
RC   STRAIN=LANGELAND;
RX   MEDLINE=98404102; PubMed=9732534;
RA   East A.K., Bhandari M., Hielm S., Collins M.D.;
RT   "Analysis of the botulinum neurotoxin type F gene clusters in
RT   proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT   barati.";
RL   Curr. Microbiol. 37:262-268(1998).
DR   EMBL; X81714; CAA57358.1; -.
DR   EMBL; L35496; AAA23210.1; -.
DR   EMBL; X70821; CAA50152.1; -.
DR   EMBL; X99064; CAA67512.1; -.
DR   PIR; S48110; S48110.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
```

```
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin.
SQ   SEQUENCE   1278 AA;  147073 MW;  A1BE1318431D6918 CRC64;

  Query Match             32.8%;  Score 61;  DB 2;  Length 1278;
  Best Local Similarity   57.1%;  Pred. No. 9.2;
  Matches    8;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              : ||::||||:|:||
Db        931 YQNFSISFWVRIPK 944


RESULT 5
Q45848
ID   Q45848      PRELIMINARY;      PRT;   361 AA.
AC   Q45848;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type B;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
DR   EMBL; X70819; CAA50150.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     361    361
SQ   SEQUENCE   361 AA;  42131 MW;  A2E0FFFC81F9533D CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 361;
  Best Local Similarity   64.3%;  Pred. No. 6;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        290 FLDFSVSFWIRIPK 303


RESULT 6
Q45846
ID   Q45846      PRELIMINARY;      PRT;   361 AA.
AC   Q45846;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type B;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
DR   EMBL; X70814; CAA50145.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     361    361
SQ   SEQUENCE   361 AA;  42175 MW;  533EA98735CD98E1 CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 361;
  Best Local Similarity   64.3%;  Pred. No. 6;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        290 FLDFSVSFWIRIPK 303


RESULT 7
Q9X708
ID   Q9X708      PRELIMINARY;      PRT;   441 AA.
AC   Q9X708;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343691; PubMed=10413679;
RA   Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA   Schiavo G.;
RT   "Functional characterisation of tetanus and botulinum neurotoxins
RT   binding domains.";
RL   J. Cell Sci. 112:2715-2724(1999).
DR   EMBL; AJ242628; CAB43706.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     441    441
SQ   SEQUENCE   441 AA;  52772 MW;  721D0B468E8C95A4 CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 441;
  Best Local Similarity   64.3%;  Pred. No. 7.5;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db         81 FLDFSVSFWIRIPK 94


RESULT 8
Q9ZAJ8
ID   Q9ZAJ8      PRELIMINARY;      PRT;  1291 AA.
AC   Q9ZAJ8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
```

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13630; CAA73968.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150840 MW;  E4D3B0E46AB2E735 CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 9
Q93G71
ID   Q93G71      PRELIMINARY;      PRT;  1291 AA.
AC   Q93G71;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin type B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1436;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF295926; AAK97132.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150824 MW;  D7CA07BAE2EB8CD2 CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 10
Q933K0
ID   Q933K0      PRELIMINARY;      PRT;  1291 AA.
AC   Q933K0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type B cryptic neurotoxin.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=593, and 588;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF300466; AAL11499.1; -.
DR   EMBL; AF300465; AAL11498.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin.
SQ   SEQUENCE   1291 AA;  150843 MW;  7AC1737B0FA5A151 CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 11
Q08077
ID   Q08077      PRELIMINARY;      PRT;  1291 AA.
AC   Q08077;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BoNT/B.
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
```

```
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Eklund 17B ATCC25765;
RX   MEDLINE=94122659; PubMed=7764370;
RA   Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT   "Nucleotide sequence of the gene coding for non-proteolytic
RT   Clostridium botulinum type B neurotoxin: comparison with other
RT   clostridial neurotoxins.";
RL   Curr. Microbiol. 28:101-110(1994).
DR   EMBL; X71343; CAA50482.1; -.
DR   PIR; I40631; I40631.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150513 MW;  71BCAFE23D69FAAA CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 12
Q8GR96
ID   Q8GR96      PRELIMINARY;      PRT;  1291 AA.
AC   Q8GR96;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin.
GN   BONTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ihara H., Kohda T., Morimoto F., Tsukamoto K., Karasawa T.,
RA   Nakamura S., Mukamoto M., Kozaki S.;
RT   "Clostridium botulinum type B neurotoxin associated with infant
RT   botulism.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB084152; BAC22064.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150574 MW;  0227CAEF4F58504D CRC64;

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 25;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 13
Q9XAV1
ID   Q9XAV1      PRELIMINARY;      PRT;  430 AA.
AC   Q9XAV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alkane 1-monooxygenase (EC 1.14.15.1).
GN   ALKB.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHA0;
RX   MEDLINE=21128526; PubMed=11207749;
RA   Smits T.H.M., Roethlisberger M., Witholt B., Van Beilen J.B.;
RT   "Molecular screening for alkane hydroxylase genes in Gram-negative and
RT   Gram-positive bacteria.";
RL   Environ. Microbiol. 1:307-317(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHA0;
RA   Smits T.H.M.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ009579; CAB51045.2; -.
DR   GO; GO:0018683; F:camphor 5-monooxygenase activity; IEA.
DR   InterPro; IPR005804; FA_desat_fam.
DR   Pfam; PF00487; FA_desaturase; 1.
KW   Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   430 AA;  48337 MW;  048E950980783E86 CRC64;

  Query Match             30.6%;  Score 57;  DB 2;  Length 430;
  Best Local Similarity   35.9%;  Pred. No. 10;
  Matches   14;  Conservative    4;  Mismatches    7;  Indels   14;  Gaps    2;

Qy          7 SFWLRVPKV------SASHLEG--------PSLHWSYGL 31
              |||  :|:       || |||         |:|||  |:
Db        224 SFWSFLPRTVWFSLSSAWHLESQRLEKLGLPTLHWKNGV 262


RESULT 14
Q9X8T8
ID   Q9X8T8      PRELIMINARY;      PRT;  502 AA.
AC   Q9X8T8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative transmembrane protein.
GN   SCO3902 OR SH24.24.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
```

```
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939118; CAB42730.1; -.
DR   PIR; T36589; T36589.
DR   GO; GO:0016021; C:integral to membrane; IEA.
KW   Transmembrane; Complete proteome.
SQ   SEQUENCE   502 AA;  54795 MW;  C84F774C0A5AADBC CRC64;

  Query Match             30.6%;  Score 57;  DB 16;  Length 502;
  Best Local Similarity   50.0%;  Pred. No. 12;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         18 SHLEGPSLHWSYGLRP 33
              :| ||| :||  | :|
Db        483 AHFEGPQVHWGKGRQP 498


RESULT 15
Q45861
ID   Q45861      PRELIMINARY;      PRT;   367 AA.
AC   Q45861;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type E (Fragment).
GN   BONT/E.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type E;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
DR   EMBL; X70818; CAA50149.1; -.
DR   PIR; S48106; S48106.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     367    367
SQ   SEQUENCE   367 AA;  42902 MW;  346A610C2FF70262 CRC64;

  Query Match             30.4%;  Score 56.5;  DB 2;  Length 367;
  Best Local Similarity   22.4%;  Pred. No. 9.9;
  Matches   15;  Conservative    8;  Mismatches    7;  Indels   37;  Gaps    2;

Qy          1 FNNFTVSFWLRVP-------------------------KVSASHLE-----------GP 23
              : ||::|||:|:|                         ||| :| |
Db        297 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNEIIWTLQDNAGINQ 356

Qy         24 SLHWSYG 30
                | ::||
Db        357 KLAFNYG 363


Search completed: March 10, 2004, 09:25:33
Job time : 35.6031 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 51.1984 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	185	99.5	34	5	AAU11421	Aau11421 Synthetic
2	185	99.5	50	5	AAU11429	Aau11429 Synthetic
3	152	81.7	32	2	AAR62702	Aar62702 LHRH-cont
4	121.5	65.3	750	3	AAY92633	Aay92633 Mutant hu
5	121	65.1	537	7	ABR82481	Abr82481 Truncated
6	121	65.1	708	7	ABR82479	Abr82479 Modified
7	121	65.1	717	7	ABR82478	Abr82478 Modified
8	119	64.0	194	6	AAO30488	Aao30488 Human TNF
9	118.5	63.7	31	3	AAY92655	Aay92655 PSMpep012
10	118	63.4	158	2	AAW81336	Aaw81336 TNF30-5,
11	118	63.4	158	5	ABB07279	Abb07279 Human TNF
12	117	62.9	750	3	AAY92627	Aay92627 Mutant hu
13	117	62.9	750	3	AAY92636	Aay92636 Mutant hu
14	116	62.4	188	3	AAY84423	Aay84423 An osteop
15	116	62.4	285	6	AAO30458	Aao30458 hIL5-P2-P
16	116	62.4	287	6	AAO30459	Aao30459 hIL5.36 v
17	116	62.4	1807	4	AAB85697	Aab85697 Recombina
18	116	62.4	2028	4	AAB85698	Aab85698 Recombina
19	115.5	62.1	693	3	AAY92647	Aay92647 Mutant hu
20	115.5	62.1	693	3	AAY92648	Aay92648 Mutant hu
21	115.5	62.1	750	3	AAY92645	Aay92645 Mutant hu
22	115.5	62.1	750	3	AAY92646	Aay92646 Mutant hu
23	115	61.8	750	3	AAY92634	Aay92634 Mutant hu
24	114.5	61.6	109	4	AAB20149	Aab20149 Growth di
25	114	61.3	158	2	AAW81333	Aaw81333 TNF30-2,
26	114	61.3	158	5	ABB07278	Abb07278 Human TNF
27	114	61.3	173	3	AAY84426	Aay84426 An osteop
28	113	60.8	713	7	ABR82480	Abr82480 Modified
29	112	60.2	21	2	AAR11896	Aar11896 Immunogen
30	112	60.2	21	2	AAW06130	Aaw06130 Tetanus t
31	112	60.2	21	2	AAR88397	Aar88397 T-cell an
32	112	60.2	21	2	AAW46449	Aaw46449 Broad ran
33	112	60.2	21	2	AAW67034	Aaw67034 Tetanus t
34	112	60.2	21	2	AAW67579	Aaw67579 T-cell ep
35	112	60.2	21	2	AAW73222	Aaw73222 Tetanus t
36	112	60.2	21	3	AAY92626	Aay92626 Foreign e
37	112	60.2	21	3	AAY99876	Aay99876 Tetanus t
38	112	60.2	21	3	AAY84428	Aay84428 Amino aci
39	112	60.2	21	3	AAY49260	Aay49260 CD4+ T ce
40	112	60.2	21	3	AAB45512	Aab45512 Tetanus P
41	112	60.2	21	4	AAE11764	Aae11764 Clostridi
42	112	60.2	21	4	AAB49072	Aab49072 Tetanus t
43	112	60.2	21	4	AAB46173	Aab46173 Tetanus t
44	112	60.2	21	4	AAB68637	Aab68637 HER-2 B c
45	112	60.2	21	4	AAB61958	Aab61958 Tetanus T

ALIGNMENTS

RESULT 1
AAU11421
ID AAU11421 standard; peptide; 34 AA.
XX
AC AAU11421;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 2.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Clostridium tetani.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .21
FT /note= "Tetanus toxoid sequence (947-967 aa)"
FT Peptide 22. .25
FT /note= "Spacer peptide"
FT Peptide 26. .34
FT /note= "Gonadotrophin releasing hormone epitope"
FT Modified-site 34
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

```
PT   or its analog.
XX
PS   Claim 11; Page 7; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 34 AA;

  Query Match             99.5%;  Score 185;  DB 5;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-20;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||||||||||||||
Db          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33


RESULT 2
AAU11429
ID   AAU11429 standard; peptide; 50 AA.
XX
AC   AAU11429;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 10.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid (947-967 aa)"
FT   Peptide         38. .41
FT                   /note= "Spacer peptide"
FT   Peptide         42. .50
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   50
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 50 AA;

  Query Match             99.5%;  Score 185;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-19;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||||||||||||||
Db         17 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49


RESULT 3
AAR62702
ID   AAR62702 standard; peptide; 32 AA.
XX
AC   AAR62702;
XX
DT   25-MAR-2003  (revised)
DT   10-SEP-1995  (first entry)
XX
DE   LHRH-containing immunogenic peptide.
XX
KW   Helper T cell epitope; universal immune stimulator; invasin; hapten;
KW   vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW   androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .22
FT                   /note= "tetanus toxin helper T cell epitope"
FT   Domain          23. .32
FT                   /note= "LHRH hapten"
XX
PN   WO9425060-A1.
XX
PD   10-NOV-1994.
XX
PF   28-APR-1994;   94WO-US004832.
XX
PR   27-APR-1993;   93US-00057166.
PR   14-APR-1994;   94US-00229275.
XX
PA   (LADD/) LADD A E.
```

PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX
PI Ladd AE, Wang CY, Zamb T;
XX
DR WPI; 1994-357910/44.
XX
PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
PS Claim 8; Page 84; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasin protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasin domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing, invasin-free immunogenic peptide as above which can be used
CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC endometriosis, benign uterine tumours, recurrent functional ovarian
CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 32 AA;

Query Match 81.7%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 5.6e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

```
Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||    ||||||||
Db          3 FNNFTVSFWLRVPKVSASHLE----HWSYGLRP 31
```

RESULT 4
AAY92633
ID AAY92633 standard; protein; 750 AA.
XX
AC AAY92633;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM1.10.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 24. .38
FT /label= P2
FT /note= "foreign epitope"
FT Peptide 673. .693
FT /label= P30
FT /note= "foreign epitope"
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page; 220pp; English.
XX
CC AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC The immunogenic analogues of PSM can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms. The method is used for inducing
CC immune responses against weakly immunogenic cell-associated peptide
CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC from the wild type human PSM (AAY92619), which appears on pages 184-187
CC of the specification
XX
SQ Sequence 750 AA;

Query Match 65.3%; Score 121.5; DB 3; Length 750;
Best Local Similarity 86.2%; Pred. No. 6.5e-09;
Matches 25; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSY 29
              |||||||||||||||||||||| || |  |
Db        673 FNNFTVSFWLRVPKVSASHLE-PSSHNKY 700
```

RESULT 5
ABR82481
ID ABR82481 standard; protein; 537 AA.
XX
AC ABR82481;
XX
DT 20-NOV-2003 (first entry)
XX
DE Truncated human CEA-TT P2 and P30 epitopes.
XX
KW CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
KW APC; cytostatic; vaccine; human; tetanus toxoid; p2; p30; antigen.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .34
FT /note= "signal peptide"
FT Protein 35. .537
FT /note= "mature protein"
XX
PN WO2003059379-A2.
XX

PD 24-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-DK000031.
XX
PR 17-JAN-2002; 2002DK-00000082.
PR 17-JAN-2002; 2002US-0350047P.
XX
PA (PHAR-) PHARMEXA AS.
XX
PI Klysner S, Voldborg B;
XX
DR WPI; 2003-587260/55.
DR N-PSDB; ACF35968.
XX
PT Inducing an immune response in humans against autologous carcinoembryonic
PT antigen (CEA) comprises administering a modified CEA polypeptide, a
PT nucleic acid encoding the polypeptide, or a microorganism expressing the
PT polypeptide.
XX
PS Disclosure; Page 134-137; 140pp; English.
XX
CC The invention relates to inducing an immune response against autologous
CC carcinoembryonic antigen (CEA) in an animal, e.g. human. The method
CC involves effecting uptake and processing by antigen presenting cells
CC (APCs) in the animal of at least 1 modified CEA polypeptide or of a
CC nucleic acid encoding the modified CEA polypeptide or of a microorganism
CC or virus expressing the modified CEA polypeptide to induce a CTL response
CC and an antibody response that targets the autologous CEA. The method is
CC useful in immunizing actively against diseases characterized by cells
CC that express CEA. The present sequence represents a truncated human CEA
CC polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced
CC in its sequence
XX
SQ Sequence 537 AA;

Query Match 65.1%; Score 121; DB 7; Length 537;
Best Local Similarity 95.8%; Pred. No. 5.3e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLEGPS 24
||||||||||||||||||||||| |
Db 513 FNNFTVSFWLRVPKVSASHLEGTS 536

RESULT 6
ABR82479
ID ABR82479 standard; protein; 708 AA.
XX
AC ABR82479;
XX
DT 20-NOV-2003 (first entry)
XX
DE Modified human CEA-TT P2 and P30 epitopes.
XX
KW CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
KW APC; cytostatic; vaccine; human; tetanus toxoid; p2; p30; antigen.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .34
FT /note= "signal peptide"
FT Protein 35. .708
FT /note= "mature protein"
XX
PN WO2003059379-A2.
XX
PD 24-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-DK000031.
XX
PR 17-JAN-2002; 2002DK-00000082.
PR 17-JAN-2002; 2002US-0350047P.
XX
PA (PHAR-) PHARMEXA AS.
XX
PI Klysner S, Voldborg B;
XX
DR WPI; 2003-587260/55.
DR N-PSDB; ACF35966.
XX
PT Inducing an immune response in humans against autologous carcinoembryonic
PT antigen (CEA) comprises administering a modified CEA polypeptide, a
PT nucleic acid encoding the polypeptide, or a microorganism expressing the
PT polypeptide.
XX
PS Disclosure; Page 121-124; 140pp; English.
XX
CC The invention relates to inducing an immune response against autologous
CC carcinoembryonic antigen (CEA) in an animal, e.g. human. The method
CC involves effecting uptake and processing by antigen presenting cells
CC (APCs) in the animal of at least 1 modified CEA polypeptide or of a
CC nucleic acid encoding the modified CEA polypeptide or of a microorganism
CC or virus expressing the modified CEA polypeptide to induce a CTL response
CC and an antibody response that targets the autologous CEA. The method is
CC useful in immunizing actively against diseases characterized by cells
CC that express CEA. The present sequence represents a modified human CEA
CC polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced
CC in its sequence
XX
SQ Sequence 708 AA;

Query Match 65.1%; Score 121; DB 7; Length 708;
Best Local Similarity 88.5%; Pred. No. 7.3e-09;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLEGPSLH 26
||||||||||||||||||||||| | |
Db 642 FNNFTVSFWLRVPKVSASHLEIPQQH 667

RESULT 7
ABR82478
ID ABR82478 standard; protein; 717 AA.
XX
AC ABR82478;
XX
DT 20-NOV-2003 (first entry)
XX
DE Modified human CEA-TT P2 and P30 epitopes.
XX
KW CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
KW APC; cytostatic; vaccine; human; tetanus toxoid; p2; p30; antigen.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .34
FT /note= "signal peptide"
FT Protein 35. .717
FT /note= "mature protein"
XX
PN WO2003059379-A2.
XX
PD 24-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-DK000031.
XX
PR 17-JAN-2002; 2002DK-00000082.
PR 17-JAN-2002; 2002US-0350047P.
XX
PA (PHAR-) PHARMEXA AS.
XX
PI Klysner S, Voldborg B;
XX
DR WPI; 2003-587260/55.

```
DR   N-PSDB; ACF35964.
XX
PT   Inducing an immune response in humans against autologous carcinoembryonic
PT   antigen (CEA) comprises administering a modified CEA polypeptide, a
PT   nucleic acid encoding the polypeptide, or a microorganism expressing the
PT   polypeptide.
XX
PS   Disclosure; Page 114-117; 140pp; English.
XX
CC   The invention relates to inducing an immune response against autologous
CC   carcinoembryonic antigen (CEA) in an animal, e.g. human. The method
CC   involves effecting uptake and processing by antigen presenting cells
CC   (APCs) in the animal of at least 1 modified CEA polypeptide or of a
CC   nucleic acid encoding the modified CEA polypeptide or of a microorganism
CC   or virus expressing the modified CEA polypeptide to induce a CTL response
CC   and an antibody response that targets the autologous CEA. The method is
CC   useful in immunizing actively against diseases characterized by cells
CC   that express CEA. The present sequence represents a modified human CEA
CC   polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced
CC   in its sequence
XX
SQ   Sequence 717 AA;

  Query Match             65.1%;  Score 121;  DB 7;  Length 717;
  Best Local Similarity   95.8%;  Pred. No. 7.4e-09;
  Matches   23;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPS 24
              |||||||||||||||||||||||| |
Db        693 FNNFTVSFWLRVPKVSASHLEGTS 716


RESULT 8
AAO30488
ID   AAO30488 standard; protein; 194 AA.
XX
AC   AAO30488;
XX
DT   22-SEP-2003  (first entry)
XX
DE   Human TNFalpha variant, TNF34-P2-P30.
XX
KW   Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;
KW   tumour necrosis factor alpha; gene therapy; arthritis; human; mutant;
KW   mutein; variant; tetanus toxoid; epitope.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          2. .109
FT                   /note= "Human TNF"
FT   Region          110. .124
FT                   /note= "Tetanus toxoid P2 epitope"
FT   Region          125. .145
FT                   /note= "Tetanus toxoid P30 epitope"
FT   Region          146. .194
FT                   /note= "Human TNF"
XX
PN   WO2003042244-A2.
XX
PD   22-MAY-2003.
XX
PF   15-NOV-2002; 2002WO-DK000764.
XX
PR   16-NOV-2001; 2001DK-00001702.
PR   16-NOV-2001; 2001US-0331575P.
XX
PA   (PHAR-) PHARMEXA AS.
PA   (KLYS/) KLYSNER S.
PA   (NIEL/) NIELSEN F S.
PA   (BRAT/) BRATT T.
PA   (VOLD/) VOLDBORG B.
PA   (MOUR/) MOURITSEN S.
XX
PI   Klysner S,  Nielsen FS,  Bratt T,  Voldborg B,  Mouritsen S;
XX
DR   WPI; 2003-449558/42.
XX
PT   New immunogenic analogue of a polymeric protein, useful for preparing a
PT   composition for treating inflammatory diseases e.g. arthritis.
XX
PS   Claim 23; Page 158; 196pp; English.
XX
CC   The invention relates to immunogenic analogues of multimeric proteins
CC   such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
CC   factor alpha (TNF, TNFalpha) and methods for production of immunogenic
CC   analogues. The immunogenic analogue is useful for preparing a composition
CC   for treating inflammatory diseases, e.g., arthritis. It is also used in
CC   gene therapy. The present sequence is human TNFalpha variant protein with
CC   an inserted tetanus toxoid P2 and P30 epitopes. This sequence is used to
CC   illustrate the method of the invention
XX
SQ   Sequence 194 AA;

  Query Match             64.0%;  Score 119;  DB 6;  Length 194;
  Best Local Similarity   81.5%;  Pred. No. 3.2e-09;
  Matches   22;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHW 27
              |||||||||||||||||||||  :  |
Db        125 FNNFTVSFWLRVPKVSASHLEAEAKPW 151


RESULT 9
AAY92655
ID   AAY92655 standard; peptide; 31 AA.
XX
AC   AAY92655;
XX
DT   10-AUG-2000  (first entry)
XX
DE   PSMpep012 - P30 inserted in hPSM insertion position 10.
XX
KW   Foreign epitope; P2; prostate specific membrane antigen; vaccination;
KW   cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW   prostate cancer; cell-associated peptide antigen.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         6. .26
FT                   /label= P30
XX
PN   WO200020027-A2.
XX
PD   13-APR-2000.
XX
PF   05-OCT-1999;   99WO-DK000525.
XX
PR   05-OCT-1998;   98DK-00001261.
PR   20-OCT-1998;   98US-0105011P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Steinaa L,  Mouritsen S,  Nielsen KG,  Haaning J,  Leach D,  Dalum I;
PI   Gautam A,  Birk P,  Karlsson G;
XX
DR   WPI; 2000-349917/30.
XX
PT   Inducing immune responses to weakly immunogenic, tumor associated peptide
PT   antigens for the treatment of breast and prostate cancer.
XX
PS   Example 1; Page 118; 220pp; English.
```

```
XX
CC   AAY92650-55 are peptides designed which correspond to the P2 and P30
CC   epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
CC   amino acids in each end. The flanking amino acids correspond to the
CC   epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
CC   cell proliferation assays, but also for ELISA or other in vitro assays.
CC   The claims detail a method for inducing immune responses against weakly
CC   immunogenic cell-associated peptide antigens (PA) such as those
CC   associated with cancers (i.e. self-proteins), for example, hPSM,
CC   heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
CC   comprises effecting simultaneous presentation by antigen producing cells
CC   (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
CC   lymphocyte) group derived from the PA and/or at least 1 B-cell group
CC   derived from the cell-associated PA; and (2) at least 1 first T helper
CC   cell group which is foreign to the animal. Analogues of human PSM, human
CC   Her2 and human/murine FGF8b comprising a substantial part of all known
CC   and predicted CTL and B-cell epitopes of the respective PA and including
CC   at least one foreign T helper epitope (e.g. P2 and/or P30) are also
CC   claimed. The method is used to treat prostate, prostate/breast or breast
CC   cancer when the PA is human PSM, FGF8b and Her2, respectively
XX
SQ   Sequence 31 AA;

  Query Match             63.7%;  Score 118.5;  DB 3;  Length 31;
  Best Local Similarity   92.3%;  Pred. No. 4.6e-10;
  Matches   24;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLH 26
              |||||||||||||||||||| || |
Db          6 FNNFTVSFWLRVPKVSASHLE-PSSH 30


RESULT 10
AAW81336
ID   AAW81336 standard; protein; 158 AA.
XX
AC   AAW81336;
XX
DT   21-APR-1999  (first entry)
XX
DE   TNF30-5, a TNF-alpha analogue.
XX
KW   Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
KW   vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
KW   cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
KW   asthma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9846642-A1.
XX
PD   22-OCT-1998.
XX
PF   15-APR-1998;   98WO-DK000157.
XX
PR   15-APR-1997;   97DK-00000418.
PR   24-APR-1997;   97US-0044187P.
XX
PA   (FERR ) FARM LAB FERRING AS.
XX
PI   Jensen MR,  Mouritsen S,  Elsner H,  Dalum I;
XX
DR   WPI; 1998-594561/50.
DR   N-PSDB; AAV68425.
XX
PT   Modified human tumour necrosis factor-alpha - comprises immunodominant T
PT   cell epitope, useful in vaccines to treat or prevent TNF-associated
PT   diseases, e.g. cancer.
XX
PS   Claim 15; Page 81-82; 134pp; English.
XX
CC   The present sequence represents a modified human tumour necrosis factor-
CC   alpha (TNF-alpha) analogue. The analogues have no residual TNF activity
CC   and are immunogenic in a large proportion of the human population (by
CC   using promiscuous epitopes). The TNF-alpha analogue is able to generate,
CC   in humans, neutralizing antibodies to wild-type human TNF alpha, has at
CC   least one fragment of TNF substituted by a peptide containing an
CC   immunodominant T-cell epitope, and at least one TNF-alpha B-cell epitope.
CC   The substitution causes a significant change in the amino acid sequence
CC   of any one of the strands in the front beta-sheet, any of the connecting
CC   loops or any of the B', I or D strands in the back beta-sheet. The TNF-
CC   alpha analogues are used as vaccines for treatment or prevention of
CC   diseases associated with excessive release or activity of TNF-alpha, e.g.
CC   rheumatoid arthritis, Crohn's disease, ulcerative colitis, cancer of any
CC   sort, disseminated sclerosis, diabetes, psoriasis, osteoporosis and
CC   asthma
XX
SQ   Sequence 158 AA;

  Query Match             63.4%;  Score 118;  DB 2;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-09;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEG 22
              ||||||||||||||||||||||
Db        133 FNNFTVSFWLRVPKVSASHLEG 154


RESULT 11
ABB07279
ID   ABB07279 standard; protein; 158 AA.
XX
AC   ABB07279;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human TNF-alpha analogue TNF30-5.
XX
KW   TNF-alpha; pharmaceutical; vaccine; self-protein; tumour necrosis factor;
KW   cetylpyridinium chloride; immunisation; antiinflammatory; antirheumatic;
KW   antiarthritic; antiulcer; cytostatic; antidiabetic; antipsoriatic;
KW   antiasthmatic; immunomodulator; neuroprotective; osteopathic; human;
KW   TNF30-5.
XX
OS   Homo sapiens.
XX
PN   WO200197837-A1.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-DK000431.
XX
PR   21-JUN-2000; 2000DK-00000966.
XX
PA   (FERR ) FERRING BV.
XX
PI   Olesen OF,  Balchen T,  Bouman MHEM;
XX
DR   WPI; 2002-114542/15.
DR   N-PSDB; ABA94389.
XX
PT   Novel vaccine composition for prevention/treatment of self-protein-
PT   mediated pathology such as cancer, diabetes and asthma, comprises
PT   modified immunogenic self-protein and surfactant capable of acting as
PT   solubilizer.
XX
PS   Claim 21; Page 42-43; 55pp; English.
XX
CC   The invention provides a pharmaceutical vaccine composition (I) for the
CC   prevention or treatment of a self-protein-mediated pathology. The
CC   composition comprises at least one modified immunogenic self-protein
CC   (selected from modified TNF-alpha proteins) and a surfactant capable of
CC   acting as a solubilizer. (I) is useful for preventing or treating a self
CC   -protein-mediated pathology such as an inflammatory disease, rheumatoid
CC   arthritis, an inflammatory bowel disease (ulcerative colitis or Crohn's
```

```
CC   disease), cancer, cachexia, multiple sclerosis, diabetes, psoriasis,
CC   osteoporosis or asthma. (I) is useful for inducing autoantibodies to a
CC   self-protein such as TNF (tumour necrosis factor)-alpha in a human
CC   subject. (I) comprising cetylpyridinium chloride as a component is useful
CC   for immunisation of a human subject or for treatment of a human
CC   inflammatory disease. The present sequence represents a human TNF-alpha
CC   analogue TNF30-5
XX
SQ   Sequence 158 AA;

  Query Match             63.4%;  Score 118;  DB 5;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-09;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEG 22
              ||||||||||||||||||||||
Db        133 FNNFTVSFWLRVPKVSASHLEG 154


RESULT 12
AAY92627
ID   AAY92627 standard; protein; 750 AA.
XX
AC   AAY92627;
XX
DT   10-AUG-2000  (first entry)
XX
DE   Mutant human prostate specific membrane antigen construct, hPSM1.1.
XX
KW   Prostate specific membrane antigen; immunogenized construct; mutant;
KW   vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW   prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         17. .31
FT                   /label= P2
FT                   /note= "foreign epitope"
FT   Peptide         32. .52
FT                   /label= P30
FT                   /note= "foreign epitope"
XX
PN   WO200020027-A2.
XX
PD   13-APR-2000.
XX
PF   05-OCT-1999;   99WO-DK000525.
XX
PR   05-OCT-1998;   98DK-00001261.
PR   20-OCT-1998;   98US-0105011P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Steinaa L,  Mouritsen S,  Nielsen KG,  Haaning J,  Leach D,  Dalum I;
PI   Gautam A,  Birk P,  Karlsson G;
XX
DR   WPI; 2000-349917/30.
XX
PT   Inducing immune responses to weakly immunogenic, tumor associated peptide
PT   antigens for the treatment of breast and prostate cancer.
XX
PS   Example 1; Page; 220pp; English.
XX
CC   AAY92627-49 are mutant immunogenized human prostate specific membrane
CC   antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC   The immunogenic analogues of PSM can be used in the claimed method as an
CC   autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC   binding regions and cysteine residues involved in disulfide bonds are
CC   preserved in the immunogenized forms. The method is used for inducing
CC   immune responses against weakly immunogenic cell-associated peptide
CC   antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC   human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC   fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC   simultaneous presentation by antigen producing cells (APCs) of the
CC   animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC   group derived from the PA and/or at least 1 B-cell group derived from the
CC   cell-associated PA; and (2) at least 1 first T helper cell group which is
CC   foreign to the animal. Analogues of human PSM, human Her2 and
CC   human/murine FGF8b comprising a substantial part of all known and
CC   predicted CTL and B-cell epitopes of the respective PA and including at
CC   least one foreign T helper epitope are also claimed. The method is used
CC   to treat prostate, prostate/breast or breast cancer when the PA is human
CC   PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC   from the wild type human PSM (AAY92619), which appears on pages 184-187
CC   of the specification
XX
SQ   Sequence 750 AA;

  Query Match             62.9%;  Score 117;  DB 3;  Length 750;
  Best Local Similarity   95.7%;  Pred. No. 3e-08;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGP 23
              |||||||||||||||||||||| |
Db         32 FNNFTVSFWLRVPKVSASHLETP 54


RESULT 13
AAY92636
ID   AAY92636 standard; protein; 750 AA.
XX
AC   AAY92636;
XX
DT   10-AUG-2000  (first entry)
XX
DE   Mutant human prostate specific membrane antigen construct, hPSM1.5.
XX
KW   Prostate specific membrane antigen; immunogenized construct; mutant;
KW   vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW   prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         24. .38
FT                   /label= P2
FT                   /note= "foreign epitope"
FT   Peptide         301. .321
FT                   /label= P30
FT                   /note= "foreign epitope"
XX
PN   WO200020027-A2.
XX
PD   13-APR-2000.
XX
PF   05-OCT-1999;   99WO-DK000525.
XX
PR   05-OCT-1998;   98DK-00001261.
PR   20-OCT-1998;   98US-0105011P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Steinaa L,  Mouritsen S,  Nielsen KG,  Haaning J,  Leach D,  Dalum I;
PI   Gautam A,  Birk P,  Karlsson G;
XX
DR   WPI; 2000-349917/30.
XX
PT   Inducing immune responses to weakly immunogenic, tumor associated peptide
PT   antigens for the treatment of breast and prostate cancer.
XX
PS   Example 1; Page; 220pp; English.
XX
CC   AAY92627-49 are mutant immunogenized human prostate specific membrane
```

```
CC   antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC   The immunogenic analogues of PSM can be used in the claimed method as an
CC   autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC   binding regions and cysteine residues involved in disulfide bonds are
CC   preserved in the immunogenized forms. The method is used for inducing
CC   immune responses against weakly immunogenic cell-associated peptide
CC   antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC   human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC   fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC   simultaneous presentation by antigen producing cells (APCs) of the
CC   animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC   group derived from the PA and/or at least 1 B-cell group derived from the
CC   cell-associated PA; and (2) at least 1 first T helper cell group which is
CC   foreign to the animal. Analogues of human PSM, human Her2 and
CC   human/murine FGF8b comprising a substantial part of all known and
CC   predicted CTL and B-cell epitopes of the respective PA and including at
CC   least one foreign T helper epitope are also claimed. The method is used
CC   to treat prostate, prostate/breast or breast cancer when the PA is human
CC   PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC   from the wild type human PSM (AAY92619), which appears on pages 184-187
CC   of the specification
XX
SQ   Sequence 750 AA;

  Query Match             62.9%;  Score 117;  DB 3;  Length 750;
  Best Local Similarity   75.8%;  Pred. No. 3e-08;
  Matches   25;  Conservative    1;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              |||||||||||||||||||||  ||    | : |
Db        301 FNNFTVSFWLRVPKVSASHLE--SLKVPYNVGP 331


RESULT 14
AAY84423
ID   AAY84423 standard; protein; 188 AA.
XX
AC   AAY84423;
XX
DT   25-JUL-2000  (first entry)
XX
DE   An osteoprotegerin ligand/tetanus toxoid P30 epitope fusion protein.
XX
KW   Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
KW   tumour necrosis factor receptor; type II transmembrane protein;
KW   osteoclast differentiation; CSF-1; osteoclast activator; immune response;
KW   osteoporosis; bone resorption.
XX
OS   Synthetic.
OS   Clostridium tetani.
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .14
FT                   /note= "His tag"
FT   Protein         15. .112
FT                   /note= "residues 158-255 of murine OPGL"
FT   Peptide         113. .133
FT                   /note= "tetanus toxoid P30 epitope"
FT   Protein         134. .188
FT                   /note= "residues 262-316 of murine OPGL"
XX
PN   WO200015807-A1.
XX
PD   23-MAR-2000.
XX
PF   13-SEP-1999;   99WO-DK000481.
XX
PR   15-SEP-1998;   98DK-00001164.
PR   02-OCT-1998;   98US-0102896P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Halkier T,  Haaning J;
XX
DR   WPI; 2000-271444/23.
DR   N-PSDB; AAZ99970.
XX
PT   In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used to
PT   treat, prevent and ameliorate osteoporosis.
XX
PS   Example; Page 94-95; 110pp; English.
XX
CC   The present sequence represents fusion protein of murine osteoprotegerin
CC   ligand (OPGL) and tetanus toxoid P30 epitope. Osteoprotegerin is a
CC   secreted member of the tumour necrosis factor receptor family, which
CC   blocks osteoclastogenesis in a dose dependent manner. The OPGL protein is
CC   synthesised as a type II transmembrane protein. The murine and human OPGL
CC   polypeptides are 87% homologous. OPGL is a potent osteoclast
CC   differentiation factor when combined with CSF-1. It is not capable of
CC   inducing osteoclast differentiation in the absence of CSF-1. OPGL is also
CC   an activator of mature osteoclasts. The specification describes a method
CC   for the in vivo down-regulation of OPGL activity in an animal. The method
CC   comprises using at least one OPGL polypeptide or subsequence, and/or at
CC   least one OPGL analogue to induce an immune response in the animal. The
CC   method and OPGL polypeptide are useful for treating, preventing and
CC   ameliorating osteoporosis or other diseases or conditions characterised
CC   by excessive bone resorption
XX
SQ   Sequence 188 AA;

  Query Match             62.4%;  Score 116;  DB 3;  Length 188;
  Best Local Similarity   82.1%;  Pred. No. 8.6e-09;
  Matches   23;  Conservative    1;  Mismatches    0;  Indels    4;  Gaps    1;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWS 28
              |||||||||||||||||||||    :||
Db        113 FNNFTVSFWLRVPKVSASHLE----NWS 136


RESULT 15
AAO30458
ID   AAO30458 standard; protein; 285 AA.
XX
AC   AAO30458;
XX
DT   22-SEP-2003  (first entry)
XX
DE   hIL5-P2-P30-hIL5 (hIL5.35) fusion construct protein.
XX
KW   Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;
KW   tumour necrosis factor alpha; gene therapy; arthritis; interleukin 5;
KW   IL5; epitope; human; tetanus toxoid; chimeric.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .19
FT                   /note= "Human IL5 leader peptide"
FT   Protein         20. .285
FT                   /note= "Mature hIL5.35 protein"
XX
PN   WO2003042244-A2.
XX
PD   22-MAY-2003.
XX
PF   15-NOV-2002; 2002WO-DK000764.
XX
PR   16-NOV-2001; 2001DK-00001702.
PR   16-NOV-2001; 2001US-0331575P.
XX
PA   (PHAR-) PHARMEXA AS.
PA   (KLYS/) KLYSNER S.
PA   (NIEL/) NIELSEN F S.
```

PA (BRAT/) BRATT T.
PA (VOLD/) VOLDBORG B.
PA (MOUR/) MOURITSEN S.
XX
PI Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;
XX
DR WPI; 2003-449558/42.
DR N-PSDB; AAL61294.
XX
PT New immunogenic analogue of a polymeric protein, useful for preparing a
PT composition for treating inflammatory diseases e.g. arthritis.
XX
PS Claim 20; Page 112-113; 196pp; English.
XX
CC The invention relates to immunogenic analogues of multimeric proteins
CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic
CC analogues. The immunogenic analogue is useful for preparing a composition
CC for treating inflammatory diseases, e.g., arthritis. It is also used in
CC gene therapy. The present sequence is a fusion construct which comprises
CC 2 human interleukin 5 (IL5) monomers joined by tetanus toxoid epitopes
CC P30 and P2. This sequence is used to illustrate the method of the
CC invention
XX
SQ Sequence 285 AA;

Query Match 62.4%; Score 116; DB 6; Length 285;
Best Local Similarity 91.7%; Pred. No. 1.4e-08;
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 FNNFTVSFWLRVPKVSASHLEGPS 24
              |||||||||||||||||||||| |:
Db        150 FNNFTVSFWLRVPKVSASHLEIPT 173
```

Search completed: March 10, 2004, 09:12:10
Job time : 52.1984 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 10.5837 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	112	60.2	1315	1	BTCLTN	tentoxilysin (EC 3
2	62	33.3	1268	2	S33411	botulinum neurotox
3	61.5	33.1	92	1	RHRTG	gonadoliberin prec
4	61	32.8	366	2	S48110	neurotoxin type F
5	61	32.8	369	2	S48109	neurotoxin type F
6	61	32.8	1274	2	I40813	neurotoxin type F
7	61	32.8	1297	2	S39791	neurotoxin - Clost
8	59	31.7	1296	1	BTCLAB	bontoxilysin (EC 3
9	58	31.2	1291	1	A48940	bontoxilysin (EC 3
10	58	31.2	1291	2	I40631	non-proteolytic bo
11	57.5	30.9	90	1	RHMSG	gonadoliberin prec
12	57	30.6	502	2	T36589	probable transmemb
13	56.5	30.4	367	2	S48106	neurotoxin type E
14	56.5	30.4	1251	2	JH0256	botulinum neurotox
15	56.5	30.4	1252	2	S21178	botulinum neurotox
16	56	30.1	812	2	T01618	hypothetical prote
17	56	30.1	1296	2	I40645	botulinum neurotox
18	54.5	29.3	92	1	RHHUG	gonadoliberin prec
19	54.5	29.3	1285	2	S70582	botulinum neurotox
20	54.5	29.3	1291	2	S46431	botulinum neurotox
21	54.5	29.3	1291	2	A49777	botulinum neurotox
22	54	29.0	67	2	I78541	gonadoliberin prec
23	53.5	28.8	469	2	B37837	probable alpha-amy
24	53.5	28.8	3122	2	T17202	DNA-directed DNA p
25	53	28.5	519	2	S78196	probable maturase
26	52	28.0	10	1	RHPGG	gonadoliberin - pi
27	52	28.0	10	1	RHSHG	gonadoliberin - sh
28	52	28.0	89	2	I51423	gonadoliberin prec
29	52	28.0	449	2	S23158	nucleocapsid prote
30	52	28.0	464	1	MNVUWC	nonstructural prot
31	52	28.0	467	1	MNVUW1	nonstructural prot
32	51.5	27.7	91	2	JC7393	medaka-type gonado
33	51.5	27.7	537	2	S78195	cytochrome-c oxida
34	51.5	27.7	573	2	S78197	probable maturase
35	51	27.4	292	2	S22441	protein kinase (EC
36	51	27.4	496	2	T38197	probable myb-like
37	51	27.4	659	2	F70453	aconitase - Aquife
38	51	27.4	944	2	T18627	hypothetical prote
39	51	27.4	1196	2	JQ1467	toxin, nontoxic co
40	51	27.4	1196	2	S46430	botulinum neurotox
41	50.5	27.2	438	2	T52082	alpha galactosyltr
42	50.5	27.2	1450	2	T30273	hypothetical prote
43	50	26.9	432	2	T04726	hypothetical prote
44	50	26.9	591	2	S04401	spheroidene monoox
45	50	26.9	836	2	JE0248	ATP-binding casset

ALIGNMENTS

RESULT 1
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with bo
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A;Experimental source: strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation; epitope region
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A;Reference number: S69348; MUID:95262688; PMID:7744050
A;Accession: S69348
A;Molecule type: protein
A;Residues: 2-31 <DEF>
C;Comment: The source of this protein was an extrachromosomal plasmid.
C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglic
C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C;Function:
A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-457/Product: tentoxylysin light chain (fragment A) #status predicted <TTL>
F;461-1315/Product: tentoxylysin heavy chain (fragment B.C) #status experimental <TTH>
F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F;233,237/Binding site: zinc (His) #status predicted
F;234/Active site: Glu #status predicted

Query Match 60.2%; Score 112; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
||||||||||||||||||||||
Db 947 FNNFTVSFWLRVPKVSASHLE 967

RESULT 2
S33411
botulinum neurotoxin type F - Clostridium barati
C;Species: Clostridium barati
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S33411; S31860
R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
FEMS Microbiol. Lett. 108, 175-182, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
A;Reference number: S33411; MUID:93252228; PMID:8486245
A;Accession: S33411
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1268 <THO>
A;Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 33.3%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 2.2;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||:||||:|:||
Db 922 YQNFSVSFWVRIPK 935

RESULT 3
RHRTG
gonadoliberin precursor - rat
N;Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormon
N;Contains: gonadoliberin; prolactin release-inhibiting factor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C;Accession: A40147; B26173; A48410
R;Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
A;Reference number: A40147; MUID:89384661; PMID:2476669
A;Accession: A40147
A;Molecule type: DNA
A;Residues: 1-92 <BON>
A;Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonado
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: B26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R;Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormon
A;Reference number: A48410; MUID:93105480; PMID:1468115
A;Accession: A48410
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-92 <MAI>
A;Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A;Experimental source: thymus
A;Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)
C;Genetics:
A;Introns: 47/3; 79/3
C;Function:
A;Description: stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activiy
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-92/Product: progonadoliberin #status predicted <PGN>
F;24-33/Product: gonadoliberin #status predicted <GLN>
F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gl

Query Match 33.1%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 0.14;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

Qy 12 VPKVSAS------HLEG-PSLHWSYGLRP 33
:||: |: ||| | ||||||||
Db 4 IPKLMAAVVLLTVCLEGCSSQHWSYGLRP 32

RESULT 4
S48110
neurotoxin type F - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S48110
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific i
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48110
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-366 <CAM>
A;Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 366;
Best Local Similarity 57.1%; Pred. No. 0.76;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||::|||:|:||
Db 297 YQNFSISFWVRIPK 310

RESULT 5

S48109
neurotoxin type F - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C;Accession: S48109
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48109
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-369 <CAM>
A;Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin

  Query Match 32.8%; Score 61; DB 2; Length 369;
  Best Local Similarity 57.1%; Pred. No. 0.77;
  Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||::|||:|:||
Db 297 YQNFSISFWVRIPK 310

RESULT 6
I40813
neurotoxin type F - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40813; S48108
R;East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E
FEMS Microbiol. Lett. 96, 225-230, 1992
A;Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A;Reference number: I40644
A;Accession: I40813
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1274 <RES>
A;Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48108
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 634-1002 <CAM>
A;Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match 32.8%; Score 61; DB 2; Length 1274;
  Best Local Similarity 57.1%; Pred. No. 3;
  Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: ||::|||:|:||
Db 930 YQNFSISFWVRIPK 943

RESULT 7
S39791
neurotoxin - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C;Accession: S39791
R;Campbell, K.; Collins, M.D.; East, A.K.
Biochim. Biophys. Acta 1216, 487-491, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium a
A;Reference number: S39791; MUID:94092745; PMID:8268233
A;Accession: S39791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1297 <CAM>
A;Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match 32.8%; Score 61; DB 2; Length 1297;
  Best Local Similarity 38.1%; Pred. No. 3.1;
  Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|:||:::||:| || : : ::
Db 930 FDNFSINFWVRTPKYNNNDIQ 950

RESULT 8
BTCLAB
bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type A
C;Species: Clostridium botulinum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with othe
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: A35294
A;Molecule type: DNA
A;Residues: 1-1296 <BIN>
A;Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
A;Experimental source: strain 62A, subtype A
R;Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990
A;Title: The complete amino acid sequence of the Clostridium botulinum type A neurotox
A;Reference number: S09492; MUID:90235864; PMID:2185020
A;Accession: S09492
A;Molecule type: DNA
A;Residues: 1,'Q',3-26,'V',28-1296 <THO>
A;Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382
A;Experimental source: NCTC 2916
R;Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin componen
A;Reference number: S67988; MUID:96096783; PMID:8521962
A;Accession: S68220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <FUJ>
A;Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
R;Betley, M.J.; Somers, E.; DasGupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A;Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-te
A;Reference number: A33401; MUID:89350959; PMID:2669749
A;Accession: A33401
A;Molecule type: DNA
A;Residues: 1-35 <BET>
A;Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
R;Gimenez, J.A.; DasGupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A;Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
A;Reference number: A53884; MUID:94000342; PMID:8397793
A;Accession: A53884
A;Status: preliminary
A;Molecule type: protein
A;Residues: 867-880;1148-1217,'Y',1219 <GIM>
A;Experimental source: strain Hall
A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after
R;DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and ar
A;Reference number: A60025; MUID:91120847; PMID:2126206
A;Accession: A60025

A;Molecule type: protein
A;Residues: 2-6;445-453,'X',455-457 <DAS1>
R;DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A;Reference number: A27000
A;Accession: A27000
A;Molecule type: protein
A;Residues: 2-47 <DAS2>
R;Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A;Title: Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.
A;Reference number: A49708; MUID:94124495; PMID:8294407
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
C;Genetics:
A;Gene: atx; botA
C;Function:
A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
F;2-444/Product: bontoxilysin A light chain #status experimental <LGHT>
F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
F;223,227/Binding site: zinc (His) #status predicted
F;224/Active site: Glu #status predicted

Query Match 31.7%; Score 59; DB 1; Length 1296;
Best Local Similarity 57.1%; Pred. No. 5.9;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: || : |||:|:||
Db 938 YENFSTSFWIRIPK 951

RESULT 9
A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type B (BoNT/B)
C;Species: Clostridium botulinum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the typ
A;Reference number: A48940; MUID:92384550; PMID:1514783
A;Accession: A48940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <WHE>
A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A;Experimental source: type B, Danish
A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publica
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-994 <CAM>
A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A;Experimental source: proteolytic type B, strain NCTC 7273
R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A;Description: Partial amino acid sequence of botulinum neurotoxin type B and comparisic
A;Reference number: S21575
A;Accession: S21575
A;Molecule type: DNA
A;Residues: 36-217,'G',219-224,'S',226-246 <SZA>
A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K
J. Biol. Chem. 267, 14721-14729, 1992

A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus t
A;Reference number: A42871; MUID:92340509; PMID:1634516
A;Accession: A42871
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-313,'S',315-451 <KUR>
A;Experimental source: strain Okra
A;Note: sequence extracted from NCBI backbone (NCBIP:109365)
R;DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
A;Reference number: S07155; MUID:89000987; PMID:3139097
A;Accession: S07155
A;Molecule type: protein
A;Residues: 2-29,'M',31-45 <DAS>
A;Accession: S08562
A;Molecule type: protein
A;Residues: 442-463,'R',465-467 <DA2>
R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A;Reference number: S07128; MUID:85197963; PMID:3888113
A;Accession: S07128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-16 <SCH1>
A;Accession: S08573
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-17 <SCH2>
A;Accession: S08574
A;Status: preliminary
A;Molecule type: protein
A;Residues: 442-459 <SCH3>
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteoly
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic syna
C;Genetics:
A;Gene: bont/b
C;Function:
A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

Query Match 31.2%; Score 58; DB 1; Length 1291;
Best Local Similarity 64.3%; Pred. No. 8.1;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
| :|:||||:|:||
Db 923 FLDFSVSFWIRIPK 936

RESULT 10
I40631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40631; S48103; S48104; S36015
R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulin
A;Reference number: I40631; MUID:94122659; PMID:7764370
A;Accession: I40631
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

```
A;Residues: 1-1291 <RES>
A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761,'E',763-841,'M',843,'T',845,'N',847-994 <CAM1>
A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A;Experimental source: non-proteolytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843,'T',845,'N',847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C;Comment: Botulinum neurotoxin type B in these strains may posses a capable catalytic s
C;Genetics:
A;Gene: bont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

  Query Match             31.2%;  Score 58;  DB 2;  Length 1291;
  Best Local Similarity   64.3%;  Pred. No. 8.1;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        923 FLDFSVSFWIRIPK 936


RESULT 11
RHMSG
gonadoliberin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A;Reference number: A47578; MUID:87069928; PMID:3024317
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C;Genetics:
A;Introns: 45/3; 77/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;22-31/Product: gonadoliberin #status predicted <GLB>
F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F;22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

  Query Match             30.9%;  Score 57.5;  DB 1;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 0.51;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         20 LEG-PSLHWSYGLRP 33
              |||  | ||||||||
Db         16 LEGCSSQHWSYGLRP 30


RESULT 12
T36589
probable transmembrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T36589
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21575
A;Accession: T36589
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-502 <OLI>
A;Cross-references: EMBL:AL049826; PIDN:CAB42730.1; GSPDB:GN00070; SCOEDB:SCH24.24
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH24.24
C;Superfamily: Mycobacterium tuberculosis probable membrane protein Rv0051

  Query Match             30.6%;  Score 57;  DB 2;  Length 502;
  Best Local Similarity   50.0%;  Pred. No. 4;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         18 SHLEGPSLHWSYGLRP 33
              :| ||| :||  | :|
Db        483 AHFEGPQVHWGKGRQP 498


RESULT 13
S48106
neurotoxin type E - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S48106
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48106
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-367 <CAM>
A;Cross-references: EMBL:X70818; NID:g407784; PIDN:CAA50149.1; PID:g407785
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match             30.4%;  Score 56.5;  DB 2;  Length 367;
  Best Local Similarity   22.4%;  Pred. No. 3.3;
  Matches   15;  Conservative    8;  Mismatches    7;  Indels   37;  Gaps    2;

Qy          1 FNNFTVSFWLRVP-------------------------KVSASHLE----------GP 23
              : ||:|||:|:|                         ||| :| |
Db        297 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNEIIWTLQDNAGINQ 356

Qy         24 SLHWSYG 30
              | ::||
Db        357 KLAFNYG 363


RESULT 14
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C;Accession: JH0256; S16145
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum ty
A;Reference number: JH0256; MUID:92181428; PMID:1543481
```

A;Accession: JH0256
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-27,'E',29-1251 <POU>
A;Cross-references: EMBL:X62088; NID:g40379
A;Experimental source: strains ATCC 43181 and ATCC 43755
R;Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.
J. Gen. Microbiol. 137, 519-525, 1991
A;Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxi
A;Reference number: S16145; MUID:91237316; PMID:2033376
A;Accession: S16145
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-229,'M',231-252 <FUJ>
A;Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408
A;Experimental source: strain BL6340
C;Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F;423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F;412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1251;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

```
Qy          1 FNNFTVSFWLRVP-------------------------KVSASHLE-----------GP 23
              : ||::|||:|:|                         |||  :| |
Db        912 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNEIIWTLQDNSGINQ 971

Qy         24 SLHWSYG 30
                | ::||
Db        972 KLAFNYG 978
```

RESULT 15
S21178
botulinum neurotoxin type E precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
C;Accession: S21178; S48107; JH0257; B35294; A60027; S18111
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992
A;Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
A;Reference number: S21178; MUID:92174922; PMID:1541280
A;Accession: S21178
A;Molecule type: DNA
A;Residues: 1-1252 <WHE>
A;Cross-references: EMBL:X62683; NID:g40397; PIDN:CAA44558.1; PID:g40398
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48107
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 616-982 <CAM>
A;Cross-references: EMBL:X70815; NID:g407786; PIDN:CAA50146.1; PID:g407787
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
R;Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinal neurotoxin E derived from Clostridium botulinum type
A;Reference number: JH0256; MUID:92181428; PMID:1543481
A;Accession: JH0257
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-176,'R',178-197,'C',199-339,'R',341-772,'I',774-962,'FE',965-966,'R',968-1
A;Cross-references: EMBL:X62089; NID:g40393; PIDN:CAA43999.1; PID:g40394
A;Experimental source: strain Beluga
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: B35294
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-176,'R',178-252 <BIN>
A;Experimental source: strain Beluga
R;Gimenez, J.A.; DasGupta, B.R.
Biochimie 72, 213-217, 1990
A;Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the s
A;Reference number: A60027; MUID:90344918; PMID:2116911
A;Accession: A60027
A;Molecule type: protein
A;Residues: 420-427 <GIM>
A;Experimental source: strain Beluga
A;Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin
C;Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit ne
C;Comment: The heavy chain mediates the binding of toxin to cell receptors while the li
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin
F;1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F;423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
F;412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1252;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

```
Qy          1 FNNFTVSFWLRVP-------------------------KVSASHLE-----------GP 23
              : ||::|||:|:|                         |||  :| |
Db        912 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNEIIWTLQDNAGINQ 971

Qy         24 SLHWSYG 30
                | ::||
Db        972 KLAFNYG 978
```

Search completed: March 10, 2004, 09:16:46
Job time : 11.6425 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.2179 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	112	60.2	1314	1	TETX_CLOTE	P04958	clostridium
2	63	33.9	431	1	PURA_LEGPN	Q8rnm2	legionella
3	61.5	33.1	92	1	GON1_RAT	P07490	rattus norv
4	61	32.8	1274	1	BXF_CLOBO	P30996	clostridium
5	61	32.8	1296	1	BXG_CLOBO	Q60393	clostridium
6	59	31.7	1295	1	BXA1_CLOBO	P10845	clostridium
7	58.5	31.5	1250	1	BXE_CLOBO	Q00496	clostridium
8	58	31.2	1290	1	BXB_CLOBO	P10844	clostridium
9	57.5	30.9	90	1	GON1_MOUSE	P13562	mus musculu
10	57	30.6	1051	1	VP2_AHSV6	O71024	african hor
11	56.5	30.4	90	1	GON1_RANCA	Q90y63	rana catesb
12	56.5	30.4	1250	1	BXE_CLOBU	P30995	clostridium
13	56	30.1	1295	1	BXA2_CLOBO	Q45894	clostridium
14	54.5	29.3	92	1	GON1_HUMAN	P01148	homo sapien
15	54.5	29.3	1290	1	BXC1_CLOBO	P18640	clostridium
16	54	29.0	67	1	GON1_MACMU	P55247	macaca mula
17	54	29.0	91	1	GON1_PIG	P49921	sus scrofa
18	53.5	28.8	760	1	AMY_CLOAB	P23671	clostridium
19	53.5	28.8	3122	1	DPOZ_MOUSE	Q61493	mus musculu
20	52	28.0	61	1	GON1_SHEEP	Q28588	ovis aries
21	52	28.0	63	1	GON1_MESAU	O09163	mesocricetu
22	52	28.0	89	1	GON1_XENLA	P45656	xenopus lae
23	52	28.0	92	1	GON1_TUPGB	Q95335	tupaia glis
24	52	28.0	449	1	VNSS_INSVN	Q01811	impatiens n
25	52	28.0	464	1	VNSS_TSWV1	P26002	tomato spot
26	52	28.0	467	1	VNSS_TSWVL	P26003	tomato spot
27	51.5	27.7	90	1	GON8_RANDY	Q9iau2	rana dybows
28	51.5	27.7	91	1	GON1_ORYLA	Q9dgc8	o progonado
29	51.5	27.7	521	1	UBP3_HUMAN	Q9y6i4	homo sapien
30	51	27.4	95	1	GON1_MORSA	O73812	morone saxa
31	51	27.4	265	1	DDHC_RHOSU	Q8gpg1	rhodovulum
32	51	27.4	292	1	CC22_ORYSA	P29619	oryza sativ
33	51	27.4	1196	1	BXCN_CLOBO	P46081	clostridium
34	50.5	27.2	537	1	COX1_SCHPO	P07657	schizosacch
35	50.5	27.2	1450	1	SREJ_STRPU	Q26627	strongyloce
36	50	26.9	241	1	CRTA_RHOCA	P17055	rhodobacter
37	50	26.9	842	1	ABC6_HUMAN	Q9np58	homo sapien
38	49.5	26.6	95	1	GON1_PAGMA	P70074	pagrus majo
39	49.5	26.6	95	1	GON1_SPAAU	P51919	sparus aura
40	49.5	26.6	3130	1	DPOZ_HUMAN	O60673	homo sapien
41	49	26.3	523	1	C9B1_GLYEC	P93149	glycyrrhiza
42	49	26.3	551	1	SYK_BRUME	Q8ycm8	brucella me
43	49	26.3	551	1	SYK_BRUSU	P59225	brucella su
44	48.5	26.1	92	1	GON1_CAVPO	O54713	cavia porce
45	48.5	26.1	5217	1	HTS1_COCCA	Q01886	cochliobolu

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=pE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RN [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

```
RL   Eur. J. Biochem. 188:39-45(1990).
RN   [6]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=92037649; PubMed=1935979;
RA   Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT   "Limited proteolysis of tetanus toxin. Relation to activity and
RT   identification of cleavage sites.";
RL   Eur. J. Biochem. 202:41-51(1991).
RN   [7]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93010948; PubMed=1396558;
RA   Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA   Montecucco C.;
RT   "Tetanus toxin is a zinc protein and its inhibition of
RT   neurotransmitter release and protease activity depend on zinc.";
RL   EMBO J. 11:3577-3583(1992).
RN   [8]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX   MEDLINE=97475217; PubMed=9334741;
RA   Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA   Sax M.;
RT   "Structure of the receptor binding fragment HC of tetanus
RT   neurotoxin.";
RL   Nat. Struct. Biol. 4:788-792(1997).
CC   -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC       BOND OF SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
CC       synaptobrevin 2.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC       YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC       AND ARE NON-TOXIC AFTER SEPARATION.
CC   -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC       GANGLIOSIDE RECEPTORS.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04436; CAA28033.1; -.
DR   EMBL; X06214; CAA29564.1; -.
DR   EMBL; AF528097; AAO37454.1; -.
DR   EMBL; M12739; AAA23282.1; -.
DR   PIR; A25689; BTCLTN.
DR   PDB; 1AF9; 29-APR-98.
DR   PDB; 1A8D; 14-OCT-98.
DR   PDB; 1D0H; 27-MAR-00.
DR   PDB; 1DFQ; 24-MAR-00.
DR   PDB; 1DIW; 24-MAR-00.
DR   PDB; 1DLL; 24-MAR-00.
DR   PDB; 1FV3; 05-SEP-01.
DR   MEROPS; M27.001; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW   3D-structure; Complete proteome.
FT   INIT_MET      0      0
FT   CHAIN         1    456       TETANUS TOXIN LIGHT CHAIN.
FT   CHAIN       457   1314       TETANUS TOXIN HEAVY CHAIN.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    233    233       BY SIMILARITY.
FT   METAL       236    236       ZINC (CATALYTIC) (BY SIMILARITY).
FT   TRANSMEM    226    246       POTENTIAL.
FT   TRANSMEM    669    689       POTENTIAL.
FT   DISULFID    438    466       INTERCHAIN.
FT   DISULFID   1076   1092
FT   HELIX       876    882
FT   TURN        883    883
FT   STRAND      884    891
FT   TURN        892    893
FT   STRAND      894    897
FT   STRAND      904    907
FT   TURN        909    910
FT   STRAND      912    915
FT   STRAND      920    925
FT   TURN        928    929
FT   STRAND      932    935
FT   HELIX       938    940
FT   TURN        941    946
FT   STRAND      949    956
FT   HELIX       962    968
FT   TURN        969    970
FT   STRAND      972    977
FT   STRAND      980    981
FT   HELIX       983    985
FT   STRAND      987    995
FT   TURN        996    997
FT   STRAND      998   1004
FT   TURN       1006   1007
FT   STRAND     1010   1016
FT   STRAND     1020   1020
FT   TURN       1021   1022
FT   STRAND     1031   1037
FT   TURN       1039   1040
FT   STRAND     1042   1047
FT   TURN       1048   1049
FT   STRAND     1050   1056
FT   TURN       1058   1059
FT   STRAND     1068   1074
FT   TURN       1079   1080
FT   STRAND     1082   1091
FT   HELIX      1097   1105
FT   TURN       1106   1107
FT   STRAND     1112   1112
FT   STRAND     1114   1114
FT   TURN       1116   1117
FT   STRAND     1120   1120
FT   STRAND     1122   1122
FT   TURN       1123   1124
FT   STRAND     1127   1131
FT   HELIX      1132   1134
FT   TURN       1135   1136
FT   STRAND     1137   1141
FT   TURN       1144   1145
FT   STRAND     1148   1152
FT   STRAND     1155   1158
FT   TURN       1159   1162
FT   STRAND     1163   1166
FT   STRAND     1173   1178
FT   TURN       1184   1185
FT   STRAND     1188   1188
FT   STRAND     1190   1190
```

```
  Query Match             60.2%;  Score 112;  DB 1;  Length 1314;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |||||||||||||||||||||
Db        946 FNNFTVSFWLRVPKVSASHLE 966


RESULT 2
PURA_LEGPN
ID   PURA_LEGPN     STANDARD;      PRT;   431 AA.
AC   Q8RNM2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE   (AdSS) (AMPSase).
GN   PURA.
OS   Legionella pneumophila.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Legionellaceae; Legionella.
OX   NCBI_TaxID=446;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rankin S., Li Z., Isberg R.R.;
RT   "Macrophage induced genes of Legionella pneumophila: protection from
RT   reactive intermediates and solute imbalance during intracellular
RT   growth.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Plays an important role in the de novo pathway of purine
CC       nucleotide biosynthesis.
CC   -!- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
CC       adenylosuccinate.
CC   -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC   -!- PATHWAY: AMP biosynthesis; first committed step.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF480918; AAM00648.1; -.
DR   HAMAP; MF_00011; -; 1.
DR   InterPro; IPR001114; Asucc_synthtase.
DR   Pfam; PF00709; Adenylsucc_synt; 1.
DR   ProDom; PD001188; Asucc_synthtase; 1.
DR   TIGRFAMs; TIGR00184; purA; 1.
DR   PROSITE; PS01266; ADENYLOSUCCIN_SYN_1; 1.
DR   PROSITE; PS00513; ADENYLOSUCCIN_SYN_2; 1.
KW   Purine biosynthesis; Ligase; GTP-binding; Metal-binding; Magnesium.
FT   NP_BIND      13     19       GTP (POTENTIAL).
FT   ACT_SITE    141    141       BY SIMILARITY.
FT   ACT_SITE    148    148       BY SIMILARITY.
FT   METAL        14     14       MAGNESIUM (BY SIMILARITY).
FT   METAL        41     41       MAGNESIUM (VIA CARBONYL OXYGEN) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   431 AA;  47381 MW;  D446C19AED550774 CRC64;

  Query Match             33.9%;  Score 63;  DB 1;  Length 431;
  Best Local Similarity   36.4%;  Pred. No. 0.25;
  Matches   12;  Conservative    7;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
              ::|| :: : : | |     | | || |:  |||
Db        168 YHNFVLTQYFKQPAVDLESLLGESLQWAEELRP 200


RESULT 3
GON1_RAT
ID   GON1_RAT       STANDARD;      PRT;    92 AA.
AC   P07490;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE   I].
GN   GNRH1 OR GNRH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86094338; PubMed=2867548;
RA   Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT   "Isolation of the gene and hypothalamic cDNA for the common precursor
RT   of gonadotropin-releasing hormone and prolactin release-inhibiting
RT   factor in human and rat.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89384661; PubMed=2476669;
RA   Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT   "The rat gonadotropin-releasing hormone: SH locus: structure and
RT   hypothalamic expression.";
RL   Mol. Endocrinol. 3:1257-1262(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=93105480; PubMed=1468115;
RA   Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT   "Thymocytes express a mRNA that is identical to hypothalamic
RT   luteinizing hormone-releasing hormone mRNA.";
RL   Cell. Mol. Neurobiol. 12:447-454(1992).
RN   [4]
RP   SEQUENCE OF 1-47 FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=87149087; PubMed=3547652;
RA   Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT   "Two mammalian genes transcribed from opposite strands of the same
RT   DNA locus.";
RL   Science 235:1514-1517(1987).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Central nervous system.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S50870; AAB24572.1; -.
DR   EMBL; M12579; AAA41263.1; -.
DR   EMBL; M31670; AAA41264.1; -.
DR   EMBL; M15527; AAA42141.1; ALT_SEQ.
DR   EMBL; M15529; AAA42139.1; -.
DR   EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR   PIR; A40147; RHRTG.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
```

```
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL       1     23
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       PROLACTIN RELEASE-INHIBITING FACTOR I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   92 AA;  10500 MW;  494B5C64DA8A3EB3 CRC64;

  Query Match             33.1%;  Score 61.5;  DB 1;  Length 92;
  Best Local Similarity   51.7%;  Pred. No. 0.076;
  Matches   15;  Conservative    3;  Mismatches    4;  Indels    7;  Gaps    2;

Qy         12 VPKVSAS------HLEG-PSLHWSYGLRP 33
              :||: |:      ||| | ||||||||
Db          4 IPKLMAAVVLLTVCLEGCSSQHWSYGLRP 32


RESULT 4
BXF_CLOBO
ID   BXF_CLOBO      STANDARD;      PRT;  1274 AA.
AC   P30996;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE   (Bontoxilysin F).
GN   BOTF.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23387;
RX   MEDLINE=93012902; PubMed=1398040;
RA   East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA   Roberts T.A., Thompson D.E.;
RT   "Sequence of the gene encoding type F neurotoxin of Clostridium
RT   botulinum.";
RL   FEMS Microbiol. Lett. 75:225-230(1992).
RN   [2]
RP   SEQUENCE OF 1-64 FROM N.A.
RC   STRAIN=Hobbs FT10;
RX   MEDLINE=94297488; PubMed=7764998;
RA   East A.K., Collins M.D.;
RT   "Conserved structure of genes encoding components of botulinum
RT   neurotoxin complex M and the sequence of the gene coding for the
RT   nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL   Curr. Microbiol. 29:69-77(1994).
RN   [3]
RP   SEQUENCE OF 634-1002 FROM N.A.
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
CC       BOND OF SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92906; AAA23263.1; -.
DR   EMBL; S73676; AAC60475.1; -.
DR   EMBL; X70820; CAA50151.1; -.
DR   EMBL; X70816; CAA50147.1; -.
DR   PIR; I40813; I40813.
DR   PIR; S48109; S48109.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    436       BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT   CHAIN       437   1274       BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT   METAL       227    227       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    228    228       BY SIMILARITY.
FT   METAL       231    231       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    445       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1274 AA;  146709 MW;  5B99756A7438B921 CRC64;

  Query Match             32.8%;  Score 61;  DB 1;  Length 1274;
  Best Local Similarity   57.1%;  Pred. No. 1.5;
  Matches    8;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              : ||::|||:|:||
Db        930 YQNFSISFWVRIPK 943


RESULT 5
BXG_CLOBO
ID   BXG_CLOBO      STANDARD;      PRT;  1296 AA.
AC   Q60393;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
DE   (Bontoxilysin G).
GN   BOTG.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
```

```
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=113 / 30;
RX   MEDLINE=94092745; PubMed=8268233;
RA   Campbell K., Collins M.D., East A.K.;
RT   "Nucleotide sequence of the gene coding for Clostridium botulinum
RT   (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT   with other clostridial neurotoxins.";
RL   Biochim. Biophys. Acta 1216:487-491(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74162; CAA52275.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    441       BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT   CHAIN       442   1296       BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    435    449       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1296 AA;  149013 MW;  DC8E47E15F665C31 CRC64;

  Query Match             32.8%;  Score 61;  DB 1;  Length 1296;
  Best Local Similarity   38.1%;  Pred. No. 1.6;
  Matches    8;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |:||:::||:| || : : ::
Db        929 FDNFSINFWVRTPKYNNNDIQ 949


RESULT 6
BXA1_CLOBO
ID   BXA1_CLOBO     STANDARD;      PRT;  1295 AA.
AC   P10845; P01561; P18639;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 2916;
RX   MEDLINE=90235864; PubMed=2185020;
RA   Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
RA   Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT   "The complete amino acid sequence of the Clostridium botulinum type A
RT   neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT   gene.";
RL   Eur. J. Biochem. 189:73-81(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=62A;
RX   MEDLINE=90264400; PubMed=2160960;
RA   Binz B., Kuarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
RT   "The complete sequence of botulinum neurotoxin type A and comparison
RT   with other clostridial neurotoxins.";
RL   J. Biol. Chem. 265:9153-9158(1990).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=62A;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN   [4]
RP   SEQUENCE OF 1-34 FROM N.A.
RC   STRAIN=Hall;
RX   MEDLINE=89350959; PubMed=2669749;
RA   Betley M.J., Somers E., Dasgupta B.R.;
RT   "Characterization of botulinum type A neurotoxin gene: delineation of
RT   the N-terminal encoding region.";
RL   Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN   [5]
RP   SEQUENCE OF 1-18 FROM N.A.
RC   STRAIN=Type A NIH;
RX   MEDLINE=96096783; PubMed=8521962;
RA   Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
RT   "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT   components of Clostridium botulinum type A progenitor toxins.";
RL   FEBS Lett. 376:41-44(1995).
RN   [6]
RP   SEQUENCE OF 1-16.
RX   MEDLINE=84178501; PubMed=6370252;
RA   Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequence of the heavy and light chains of
RT   botulinum neurotoxin type A.";
RL   Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN   [7]
RP   SEQUENCE OF 1-46.
RA   Dasgupta B.R., Foley J., Niece R.;
RT   "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL   Biochemistry 26:4162-4162(1987).
RN   [8]
RP   SEQUENCE OF 1-5 AND 444-456.
RX   MEDLINE=91120847; PubMed=2126206;
RA   Dasgupta B.R., Dekleva M.L.;
RT   "Botulinum neurotoxin type A: sequence of amino acids at the
RT   N-terminus and around the nicking site.";
RL   Biochimie 72:661-664(1990).
RN   [9]
```

```
RP   SEQUENCE OF 448-464 AND 872-895.
RX   MEDLINE=89024662; PubMed=3178218;
RA   Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT   "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RT   halves and their partial sequences.";
RL   Arch. Biochem. Biophys. 266:142-151(1988).
RN   [10]
RP   SEQUENCE OF 448-482.
RX   MEDLINE=85285016; PubMed=3896784;
RA   Shone C.C., Hambleton P., Melling J.;
RT   "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RT   and purification of two tryptic fragments. Proteolytic action near
RT   the COOH-terminus of the heavy subunit destroys toxin-binding
RT   activity.";
RL   Eur. J. Biochem. 151:75-82(1985).
RN   [11]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [12]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
RN   [13]
RP   MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX   MEDLINE=21556941; PubMed=11700044;
RA   Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
RT   "Site-directed mutagenesis identifies active-site residues of the
RT   light chain of botulinum neurotoxin type a.";
RL   Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX   MEDLINE=98455071; PubMed=9783750;
RA   Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT   "Crystal structure of botulinum neurotoxin type A and implications
RT   for toxicity.";
RL   Nat. Struct. Biol. 5:898-902(1998).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
CC       the treatment of strabismus and blepharospasm associated with
CC       dystonia and cervical dystonia. Also used for the treatment of
CC       hemifacial spasm and a number of other neurological disorders
CC       characterized by abnormal muscle contraction.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   -!- DATABASE: NAME=BOTOX product information Web site;
CC       WWW="http://www.botox.com/index.jsp?hp&productinfo".
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 19 of February 2002;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt019.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52066; CAA36289.1; -.
DR   EMBL; M30196; AAA23262.1; -.
DR   EMBL; X92973; CAA63551.1; -.
DR   EMBL; D67030; BAA11051.1; -.
DR   EMBL; M27892; AAA23269.1; -.
DR   PIR; A35294; BTCLAB.
DR   PDB; 3BTA; 01-OCT-99.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   Pharmaceutical; 3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC).
FT   ACT_SITE    223    223
FT   METAL       226    226       ZINC (CATALYTIC).
FT   METAL       261    261       ZINC (CATALYTIC).
FT   DISULFID    429    453       INTERCHAIN.
FT   DISULFID   1234   1279
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
FT   VARIANT      26     26       V -> A.
FT   MUTAGEN     261    261       E->A: DRASTIC DECREASE IN ENZYMATIC
FT                                ACTIVITY.
FT   MUTAGEN     265    265       F->A: DECREASES ENZYMATIC ACTIVITY.
FT   MUTAGEN     365    365       Y->A: DECREASES ENZYMATIC ACTIVITY.
FT   CONFLICT      1      1       P -> Q (IN REF. 1).
FT   CONFLICT    479    479       E -> P (IN REF. 9).
FT   CONFLICT    875    875       T -> L (IN REF. 8).
FT   CONFLICT    891    891       S -> K (IN REF. 8).
SQ   SEQUENCE   1295 AA;  149322 MW;  858342F754862579 CRC64;

  Query Match             31.7%;  Score 59;  DB 1;  Length 1295;
  Best Local Similarity   57.1%;  Pred. No. 3;
  Matches    8;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              : ||: |||:|:|| 
Db        937 YENFSTSFWIRIPK 950


RESULT 7
BXE_CLOBO
ID   BXE_CLOBO      STANDARD;      PRT;  1250 AA.
AC   Q00496;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beluga;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92174922; PubMed=1541280;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RT   "The complete amino acid sequence of the Clostridium botulinum type-E
RT   neurotoxin, derived by nucleotide-sequence analysis of the encoding
RT   gene.";
RL   Eur. J. Biochem. 204:657-667(1992).
RN   [3]
RP   SEQUENCE OF 1-251 FROM N.A.
RX   MEDLINE=90264400; PubMed=2160960;
RA   Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
RT   "The complete sequence of botulinum neurotoxin type A and comparison
RT   with other clostridial neurotoxins.";
RL   J. Biol. Chem. 265:9153-9158(1990).
RN   [4]
RP   SEQUENCE OF 1-13.
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [5]
RP   SEQUENCE OF 419-426.
RX   MEDLINE=90344918; PubMed=2116911;
RA   Gimenez J.A., Dasgupta B.R.;
RT   "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT   reveals the site trypsin nicks and homology with tetanus
RT   neurotoxin.";
RL   Biochimie 72:213-217(1990).
RN   [6]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC       181 BOND IN SNAP-25.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62089; CAA43999.1; -.
DR   EMBL; X62683; CAA44558.1; -.
DR   PIR; S08575; S08575.
DR   PIR; S21178; S21178.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    176    176       R -> G (IN REF. 2).
FT   CONFLICT    197    197       C -> S (IN REF. 2 AND 3).
FT   CONFLICT    339    339       R -> A (IN REF. 2).
FT   CONFLICT    772    772       I -> L (IN REF. 2).
FT   CONFLICT    962    963       FE -> LQ (IN REF. 2).
FT   CONFLICT    966    966       R -> A (IN REF. 2).
FT   CONFLICT   1194   1194       N -> NN (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143712 MW;  D9FCE26DDA041EB4 CRC64;

  Query Match             31.5%;  Score 58.5;  DB 1;  Length 1250;
  Best Local Similarity   23.6%;  Pred. No. 3.4;
  Matches   13;  Conservative    9;  Mismatches    4;  Indels   29;  Gaps    2;

Qy          1 FNNFTVSFWLRVP-----------------------KVSASHLEGPSLHWSY 29
              : ||::|||:|:|                       ||| :| |   : |::
Db        911 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNE---IIWTF 962


RESULT 8
BXB_CLOBO
ID   BXB_CLOBO      STANDARD;      PRT;  1290 AA.
AC   P10844; P10843;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE   (Bontoxilysin B).
GN   BOTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92384550; PubMed=1514783;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA   Minton N.P.;
RT   "Molecular cloning of the Clostridium botulinum structural gene
RT   encoding the type B neurotoxin and determination of its entire
RT   nucleotide sequence.";
RL   Appl. Environ. Microbiol. 58:2345-2354(1992).
RN   [2]
```

```
RP   SEQUENCE OF 35-245 FROM N.A.
RC   STRAIN=NCTC 7273;
RA   Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 633-993 FROM N.A.
RC   STRAIN=NCTC 7273;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-44 AND 441-466.
RC   STRAIN=657;
RX   MEDLINE=89000987; PubMed=3139097;
RA   Dasgupta B.R., Datta A.;
RT   "Botulinum neurotoxin type B (strain 657): partial sequence and
RT   similarity with tetanus toxin.";
RL   Biochimie 70:811-817(1988).
RN   [5]
RP   SEQUENCE OF 1-16 AND 441-458.
RC   STRAIN=OKRA;
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93054694; PubMed=1429690;
RA   Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT   "Botulinum neurotoxins are zinc proteins.";
RL   J. Biol. Chem. 267:23479-23483(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC       SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81186; AAA23211.1; -.
DR   EMBL; Z11934; CAA77991.1; -.
DR   EMBL; X70817; CAA50148.1; -.
DR   PIR; A48940; A48940.
DR   PDB; 1EPW; 01-NOV-00.
DR   PDB; 1F31; 01-NOV-00.
DR   PDB; 1F82; 16-AUG-00.
DR   PDB; 1F83; 16-AUG-00.
DR   PDB; 1FQH; 06-DEC-00.
DR   PDB; 1G9A; 13-NOV-02.
DR   PDB; 1G9B; 13-NOV-02.
DR   PDB; 1G9C; 13-NOV-02.
DR   PDB; 1G9D; 13-NOV-02.
DR   PDB; 1I1E; 21-NOV-01.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    440       BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT   CHAIN       441   1290       BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    445       INTERCHAIN (PROBABLE).
FT   CONFLICT     29     29       T -> M (IN REF. 4).
FT   CONFLICT    217    217       R -> G (IN REF. 2).
FT   CONFLICT    224    224       A -> S (IN REF. 2).
FT   CONFLICT    463    463       S -> R (IN REF. 4).
SQ   SEQUENCE   1290 AA;  150670 MW;  D21746E2C024DF43 CRC64;

  Query Match             31.2%;  Score 58;  DB 1;  Length 1290;
  Best Local Similarity   64.3%;  Pred. No. 4.1;
  Matches    9;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPK 14
              | :|:||||:|:||
Db        922 FLDFSVSFWIRIPK 935


RESULT 9
GON1_MOUSE
ID   GON1_MOUSE     STANDARD;      PRT;    90 AA.
AC   P13562;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE   I].
GN   GNRH1 OR GNRH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87069928; PubMed=3024317;
RA   Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA   Phillips H.S., Nikolics K., Seeburg P.H.;
RT   "A deletion truncating the gonadotropin-releasing hormone gene is
RT   responsible for hypogonadism in the hpg mouse.";
RL   Science 234:1366-1371(1986).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
```

```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14872; AAA37717.1; -.
DR   PIR; A47578; RHMSG.
DR   MGD; MGI:95789; Gnrh.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     21
FT   CHAIN        22     90       PROGONADOLIBERIN I.
FT   PEPTIDE      22     31       GONADOLIBERIN I.
FT   PEPTIDE      35     90       PROLACTIN RELEASE-INHIBITING FACTOR I.
FT   ACT_SITE     24     24       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      22     22       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      31     31       AMIDATION (G-32 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   90 AA;  10337 MW;  1C0766FA4826E4D9 CRC64;

  Query Match             30.9%;  Score 57.5;  DB 1;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 0.27;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         20 LEG-PSLHWSYGLRP 33
              |||  | ||||||||
Db         16 LEGCSSQHWSYGLRP 30


RESULT 10
VP2_AHSV6
ID   VP2_AHSV6      STANDARD;      PRT;  1051 AA.
AC   O71024;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Outer capsid protein VP2.
GN   S2 OR L2.
OS   African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS   (serotype 6)).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=86060;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98278331; PubMed=9617769;
RA   Williams C.F., Inoue T., Lucus A.-M., Zanotto P., Roy P.;
RT   "The complete sequence of four major structural proteins of African
RT   horse sickness virus serotype 6: evolutionary relationships within
RT   and between the orbiviruses.";
RL   Virus Res. 53:53-73(1998).
CC   -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC       WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC       MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC   -!- SIMILARITY: Belongs to the reoviruses VP2 protein family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021235; AAC40994.1; -.
DR   InterPro; IPR001742; Orbi_VP2.
DR   Pfam; PF00898; Orbi_VP2; 1.
KW   Coat protein.
SQ   SEQUENCE   1051 AA;  122326 MW;  2B04DB9E389F4B5F CRC64;

  Query Match             30.6%;  Score 57;  DB 1;  Length 1051;
  Best Local Similarity   47.6%;  Pred. No. 4.6;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 FNNFTVSFWLRVPKVSASHLE 21
              |:   ||:| || |::  |||
Db        636 FSKRFVSYWYRVEKITTKHLE 656


RESULT 11
GON1_RANCA
ID   GON1_RANCA     STANDARD;      PRT;    90 AA.
AC   Q90Y63;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE   (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I (GAP1)].
GN   GNRH1 OR GNRH.
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC   TISSUE=Forebrain;
RX   MEDLINE=21102951; PubMed=11170016;
RA   Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA   Kwon H.B.;
RT   "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT   precursors from bullfrog (Rana catesbeiana).";
RL   J. Exp. Zool. 289:190-201(2001).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Forebrain.
CC   -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC       during post-breeding. Not expressed in pituitary.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF188754; AAL05972.1; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
DR   GO; GO:0009755; P:hormone mediated signaling; NAS.
DR   GO; GO:0000003; P:reproduction; NAS.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25     90       PROGONADOLIBERIN I.
FT   PEPTIDE      25     34       GONADOLIBERIN I.
FT   PEPTIDE      38     86       GNRH-ASSOCIATED PEPTIDE I (BY
FT                                SIMILARITY).
FT   MOD_RES      25     25       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
```

```
FT   MOD_RES      34     34       AMIDATION (G-35 PROVIDE AMIDE GROUP) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   90 AA;  10291 MW;  317203B4E3DA2FE7 CRC64;

  Query Match             30.4%;  Score 56.5;  DB 1;  Length 90;
  Best Local Similarity   46.7%;  Pred. No. 0.38;
  Matches   14;  Conservative    4;  Mismatches    9;  Indels    3;  Gaps    2;

Qy          5 TVSFWLRVPKVSASH-LEGPSLHWSYGLRP 33
              ||   | :  : :|| : |   ||||||||
Db          6 TVVLLLAIVLLLSSHMIHGQ--HWSYGLRP 33


RESULT 12
BXE_CLOBU
ID   BXE_CLOBU      STANDARD;      PRT;  1250 AA.
AC   P30995;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
DE   (Bontoxilysin E).
OS   Clostridium butyricum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1492;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 43181, and ATCC 43755;
RX   MEDLINE=92181428; PubMed=1543481;
RA   Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT   "Sequences of the botulinal neurotoxin E derived from Clostridium
RT   botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT   ATCC 43181 and ATCC 43755).";
RL   Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN   [2]
RP   SEQUENCE OF 1-251 FROM N.A.
RC   STRAIN=BL6340;
RX   MEDLINE=91237316; PubMed=2033376;
RA   Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA   Yokosawa N., Yashiki T., Oguma K.;
RT   "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT   type E toxin gene from Clostridium butyricum strain BL6340.";
RL   J. Gen. Microbiol. 137:519-525(1991).
RN   [3]
RP   SEQUENCE OF 1-48.
RC   STRAIN=5262;
RA   Gimenez J., Foley J., Dasgupta B.R.;
RT   "Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT   partial sequence and comparison.";
RL   FASEB J. 2:A1750-A1750(1988).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62088; CAA43998.1; -.
DR   EMBL; X53180; CAA37321.1; -.
DR   PIR; JH0256; JH0256.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    421       BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT   CHAIN       422   1250       BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT   METAL       211    211       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    212    212       BY SIMILARITY.
FT   METAL       215    215       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    411    425       INTERCHAIN (PROBABLE).
FT   CONFLICT    229    229       K -> M (IN REF. 2).
SQ   SEQUENCE   1250 AA;  143265 MW;  8171B5B2C2312857 CRC64;

  Query Match             30.4%;  Score 56.5;  DB 1;  Length 1250;
  Best Local Similarity   22.4%;  Pred. No. 6.5;
  Matches   15;  Conservative    8;  Mismatches    7;  Indels   37;  Gaps    2;

Qy          1 FNNFTVSFWLRVP-------------------------KVSASHLE-----------GP 23
              : ||::|||:|:|                         ||| :| |
Db        911 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNSGWKVSLNHNEIIWTLQDNSGINQ 970

Qy         24 SLHWSYG 30
                | ::||
Db        971 KLAFNYG 977


RESULT 13
BXA2_CLOBO
ID   BXA2_CLOBO     STANDARD;      PRT;  1295 AA.
AC   Q45894; P77780;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=94143603; PubMed=8310180;
RA   Willems A., East A.K., Lawson P.A., Collins M.D.;
RT   "Sequence of the gene coding for the neurotoxin of Clostridium
RT   botulinum type A associated with infant botulism: comparison with
RT   other clostridial neurotoxins.";
RL   Res. Microbiol. 144:547-556(1993).
RN   [2]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
```

```
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure (By similarity).
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73423; CAA51824.1; -.
DR   EMBL; X87974; CAA61234.1; -.
DR   PIR; I40645; I40645.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    223    223       BY SIMILARITY.
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    453       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   1234   1279       BY SIMILARITY.
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
SQ   SEQUENCE   1295 AA;  149279 MW;  5DA04A13D98D6372 CRC64;

  Query Match             30.1%;  Score 56;  DB 1;  Length 1295;
  Best Local Similarity   50.0%;  Pred. No. 7.9;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FNNPFTVSFWLRVPK 14
              : ||: |||:::|| 
Db        937 YENPFSTSFWIKIPK 950


RESULT 14
GON1_HUMAN
ID   GON1_HUMAN     STANDARD;      PRT;    92 AA.
AC   P01148;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE   peptide I].
GN   GNRH1 OR GNRH OR LHRH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89366682; PubMed=2671939;
RA   Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT   "The complete nucleotide sequence of the human gonadotropin-releasing
RT   hormone gene.";
RL   Nucleic Acids Res. 17:6403-6403(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86094338; PubMed=2867548;
RA   Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT   "Isolation of the gene and hypothalamic cDNA for the common precursor
RT   of gonadotropin-releasing hormone and prolactin release-inhibiting
RT   factor in human and rat.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT SER-16.
RX   MEDLINE=85012739; PubMed=6090951;
RA   Seeburg P.H., Adelman J.P.;
RT   "Characterization of cDNA for precursor of human luteinizing hormone
RT   releasing hormone.";
RL   Nature 311:666-668(1984).
RN   [4]
RP   SEQUENCE OF 24-33.
RX   MEDLINE=83126573; PubMed=6760865;
RA   Tan L., Rousseau P.;
RT   "The chemical identity of the immunoreactive LHRH-like peptide
RT   biosynthesized in the human placenta.";
RL   Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN   [5]
RP   VARIANT SER-16.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [6]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC       Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
CC       (Serono).
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X01059; CAA25526.1; -.
DR   EMBL; M12578; AAA35916.1; -.
DR   EMBL; X15215; CAA33285.1; -.
DR   PIR; S05308; RHHUG.
DR   Genew; HGNC:4419; GNRH1.
DR   MIM; 152760; -.
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Pharmaceutical; Signal; Polymorphism;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL       1     23
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT   VARIANT      16     16       W -> S (in dbSNP:6185).
FT                                /FTId=VAR_013943.
SQ   SEQUENCE   92 AA;  10380 MW;  30A72221B076FA79 CRC64;

  Query Match             29.3%;  Score 54.5;  DB 1;  Length 92;
  Best Local Similarity   73.3%;  Pred. No. 0.73;
  Matches   11;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         20 LEG-PSLHWSYGLRP 33
              :|| | |||||||||
Db         18 VEGCSSQHWSYGLRP 32


RESULT 15
BXC1_CLOBO
ID   BXC1_CLOBO     STANDARD;      PRT;  1290 AA.
AC   P18640;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE   (Bontoxilysin C1).
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90370487; PubMed=2204031;
RA   Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
RA   Boquet P., Popoff M.R.;
RT   "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL   Nucleic Acids Res. 18:4924-4924(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=91024998; PubMed=2222445;
RA   Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA   Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT   "The complete nucleotide sequence of the gene coding for botulinum
RT   type C1 toxin in the C-ST phage genome.";
RL   Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN   [3]
RP   SEQUENCE OF 2-25.
RC   STRAIN=Type C Stockholm / C-ST;
RX   MEDLINE=88153072; PubMed=2450068;
RA   Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
RA   Oguma K.;
RT   "Establishment of a monoclonal antibody recognizing an antigenic site
RT   common to Clostridium botulinum type B, C1, D, and E toxins and
RT   tetanus toxin.";
RL   Infect. Immun. 56:898-902(1988).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94038966; PubMed=7901002;
RA   Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT   "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
RT   cleaving HPC-1/syntaxin.";
RL   EMBO J. 12:4821-4828(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
CC       STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC       BACTERIOPHAGE.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66433; CAA47060.1; -.
DR   EMBL; X72793; CAA51313.1; -.
DR   EMBL; X53751; CAA37780.1; -.
DR   EMBL; D90210; BAA14235.1; -.
DR   EMBL; X62389; CAA44263.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0
FT   CHAIN         1    448       BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
FT   CHAIN       449   1290       BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
FT   METAL       228    228       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    229    229       BY SIMILARITY.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    452       INTERCHAIN (PROBABLE).
FT   CONFLICT     84     84       P -> T (IN REF. 2).
SQ   SEQUENCE   1290 AA;  148734 MW;  71FBE379F97129E8 CRC64;

  Query Match             29.3%;  Score 54.5;  DB 1;  Length 1290;
  Best Local Similarity   30.8%;  Pred. No. 13;
  Matches   12;  Conservative   10;  Mismatches    8;  Indels    9;  Gaps    2;
```

```
Qy          1 FNNFTVSFWLRVPKVSASHLEGPSL--------HWSYGL 31
              : :|::|||:|: |    |:| | ::          || |:
Db        934 YESFSISFWIRINK-WVSNLPGYTIIDSVKNNSGWSIGI 971
```

Search completed: March 10, 2004, 09:13:53
Job time : 7.2179 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 12.0019 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	117	73.1	42	1	US-08-446-692-20	Sequence 20, Appl
2	117	73.1	42	2	US-08-488-351A-20	Sequence 20, Appl
3	116	72.5	27	1	US-08-446-692-19	Sequence 19, Appl
4	116	72.5	27	2	US-08-488-351A-19	Sequence 19, Appl
5	116	72.5	27	3	US-09-100-414B-36	Sequence 36, Appl
6	116	72.5	27	3	US-09-303-323-36	Sequence 36, Appl
7	116	72.5	27	4	US-09-770-014-36	Sequence 36, Appl
8	116	72.5	45	1	US-08-446-692-33	Sequence 33, Appl
9	116	72.5	45	2	US-08-488-351A-33	Sequence 33, Appl
10	113	70.6	27	3	US-09-100-414B-43	Sequence 43, Appl
11	113	70.6	27	3	US-09-303-323-43	Sequence 43, Appl
12	113	70.6	27	4	US-09-770-014-43	Sequence 43, Appl
13	110	68.8	31	3	US-09-100-414B-55	Sequence 55, Appl
14	110	68.8	31	3	US-09-303-323-55	Sequence 55, Appl
15	110	68.8	31	4	US-09-770-014-55	Sequence 55, Appl
16	108	67.5	27	3	US-09-100-414B-41	Sequence 41, Appl
17	108	67.5	27	3	US-09-100-414B-47	Sequence 47, Appl
18	108	67.5	27	3	US-09-303-323-41	Sequence 41, Appl
19	108	67.5	27	3	US-09-303-323-47	Sequence 47, Appl
20	108	67.5	27	4	US-09-770-014-41	Sequence 41, Appl
21	108	67.5	27	4	US-09-770-014-47	Sequence 47, Appl
22	108	67.5	45	3	US-09-100-414B-45	Sequence 45, Appl
23	108	67.5	45	3	US-09-303-323-45	Sequence 45, Appl
24	108	67.5	45	4	US-09-770-014-45	Sequence 45, Appl
25	107	66.9	31	3	US-09-100-414B-59	Sequence 59, Appl
26	107	66.9	31	3	US-09-303-323-59	Sequence 59, Appl
27	107	66.9	31	4	US-09-770-014-59	Sequence 59, Appl
28	106	66.2	28	3	US-09-100-414B-38	Sequence 38, Appl
29	106	66.2	28	3	US-09-303-323-38	Sequence 38, Appl
30	106	66.2	28	4	US-09-770-014-38	Sequence 38, Appl
31	105	65.6	31	3	US-09-100-414B-53	Sequence 53, Appl
32	105	65.6	31	3	US-09-303-323-53	Sequence 53, Appl
33	105	65.6	31	4	US-09-770-014-53	Sequence 53, Appl
34	105	65.6	47	3	US-09-100-414B-60	Sequence 60, Appl
35	105	65.6	47	3	US-09-303-323-60	Sequence 60, Appl
36	105	65.6	47	4	US-09-770-014-60	Sequence 60, Appl
37	105	65.6	49	3	US-09-100-414B-57	Sequence 57, Appl
38	105	65.6	49	3	US-09-303-323-57	Sequence 57, Appl
39	105	65.6	49	4	US-09-770-014-57	Sequence 57, Appl
40	103.5	64.7	25	1	US-08-446-692-17	Sequence 17, Appl
41	103.5	64.7	25	2	US-08-488-351A-17	Sequence 17, Appl
42	103	64.4	27	3	US-09-100-414B-50	Sequence 50, Appl
43	103	64.4	27	3	US-09-303-323-50	Sequence 50, Appl
44	103	64.4	27	4	US-09-770-014-50	Sequence 50, Appl
45	103	64.4	35	3	US-09-100-414B-80	Sequence 80, Appl

ALIGNMENTS

```
RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-446-692-20

  Query Match           73.1%;  Score 117;  DB 1;  Length 42;
  Best Local Similarity  82.8%;  Pred. No. 4.7e-11;
  Matches   24;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          2 LLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||||||| |   |||||||
Db         15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41
```

```
RESULT 2
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,351A
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,275
;      FILING DATE:  14-APR-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/057,166
;      FILING DATE:  27-APR-1992
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-488-351A-20

  Query Match             73.1%;  Score 117;  DB 2;  Length 42;
  Best Local Similarity   82.8%;  Pred. No. 4.7e-11;
  Matches   24;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          2 LLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||||||||| |   ||||||||
Db         15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41


RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-446-692-19

  Query Match             72.5%;  Score 116;  DB 1;  Length 27;
  Best Local Similarity   85.7%;  Pred. No. 4e-11;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||||| |   ||||||||
Db          1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26


RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,351A
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 72.5%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
||||||||||||||| | ||||||||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 5
US-09-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 72.5%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
||||||||||||||| | ||||||||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 6
US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 72.5%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 4e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
||||||||||||||| | ||||||||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 7
US-09-770-014-36
; Sequence 36, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York

```
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10154-0054
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC Windows
;       SOFTWARE:  Word 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/770,014
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/100,414
;       FILING DATE:  20-JUNE-1998
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4157
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-758-4800
;       TELEFAX:  212-751-6849
;   INFORMATION FOR SEQ ID NO:  36:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  27 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-770-014-36

  Query Match             72.5%;  Score 116;  DB 4;  Length 27;
  Best Local Similarity   85.7%;  Pred. No. 4e-11;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||| |   ||||||||
Db          1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26


RESULT 8
US-08-446-692-33
; Sequence 33, Application US/08446692
; Patent No. 5759551
;  GENERAL INFORMATION:
;     APPLICANT:  Ladd, Anna
;     APPLICANT:  Wang, Chang Yi
;     APPLICANT:  Zamb, Timothy
;     TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;     TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;     NUMBER OF SEQUENCES:  114
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Maria C.H. Lin
;       STREET:  345 Park Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  US
;       ZIP:  10154-0053
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/446,692
;       FILING DATE:  7-JUN-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria C.H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4146 US2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)415-8745
;       TELEFAX:  (516)751-6849
;   INFORMATION FOR SEQ ID NO:  33:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  45 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-446-692-33

  Query Match             72.5%;  Score 116;  DB 1;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 7.3e-11;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||| |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 9
US-08-488-351A-33
; Sequence 33, Application US/08488351A
; Patent No. 5843446
;  GENERAL INFORMATION:
;     APPLICANT:  Ladd, Anna
;     APPLICANT:  Wang, Chang Yi
;     APPLICANT:  Zamb, Timothy
;     TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;     TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;     NUMBER OF SEQUENCES:  114
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Maria C.H. Lin
;       STREET:  345 Park Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  US
;       ZIP:  10154-0053
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,351A
;       FILING DATE:  7-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/446,692
;       FILING DATE:  7-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/229,275
;       FILING DATE:  14-APR-1994
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/057,166
;       FILING DATE:  27-APR-1992
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria C.H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4146 US2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)415-8745
;       TELEFAX:  (516)751-6849
;   INFORMATION FOR SEQ ID NO:  33:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  45 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-488-351A-33

  Query Match             72.5%;  Score 116;  DB 2;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 7.3e-11;
```

```
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||||  |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 10
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-43

  Query Match             70.6%;  Score 113;  DB 3;  Length 27;
  Best Local Similarity   82.1%;  Pred. No. 1.2e-10;
  Matches   23;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:|||| |   ||||||||
Db          1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26


RESULT 11
US-09-303-323-43
; Sequence 43, Application US/09303323
; Patent No. 6228987
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/303,323
;      FILING DATE:  30-APR-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-303-323-43

  Query Match             70.6%;  Score 113;  DB 3;  Length 27;
  Best Local Similarity   82.1%;  Pred. No. 1.2e-10;
  Matches   23;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:|||| |   ||||||||
Db          1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26


RESULT 12
US-09-770-014-43
; Sequence 43, Application US/09770014
; Patent No. 6559282
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/770,014
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-770-014-43

  Query Match             70.6%;  Score 113;  DB 4;  Length 27;
  Best Local Similarity   82.1%;  Pred. No. 1.2e-10;
  Matches   23;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:|||| |   ||||||||
Db          1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26


RESULT 13
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-55

  Query Match             68.8%;  Score 110;  DB 3;  Length 31;
  Best Local Similarity   78.6%;  Pred. No. 4e-10;
  Matches   22;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:||||      ||||||||
Db          3 LSEIKGVIVHKLEGVLFGGEHWSYGLRP 30


RESULT 14
US-09-303-323-55
; Sequence 55, Application US/09303323
; Patent No. 6228987
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/303,323
;      FILING DATE:  30-APR-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-303-323-55

  Query Match             68.8%;  Score 110;  DB 3;  Length 31;
  Best Local Similarity   78.6%;  Pred. No. 4e-10;
  Matches   22;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:||||      ||||||||
Db          3 LSEIKGVIVHKLEGVLFGGEHWSYGLRP 30


RESULT 15
US-09-770-014-55
; Sequence 55, Application US/09770014
; Patent No. 6559282
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/770,014
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
```

```
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-770-014-55

  Query Match             68.8%;  Score 110;  DB 4;  Length 31;
  Best Local Similarity   78.6%;  Pred. No. 4e-10;
  Matches   22;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:||||     ||||||||
Db          3 LSEIKGVIVHKLEGVLFGGEHWSYGLRP 30


Search completed: March 10, 2004, 09:28:53
Job time : 12.0019 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 24.3658 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	159	99.4	31	9	US-09-848-834A-9	Sequence 9, Appli
2	159	99.4	47	9	US-09-848-834A-17	Sequence 17, Appl
3	116	72.5	45	14	US-10-076-674-9	Sequence 9, Appli
4	116	72.5	45	15	US-10-355-161A-9	Sequence 9, Appli
5	99	61.9	75	10	US-09-847-102A-33	Sequence 33, Appl
6	99	61.9	75	14	US-10-285-976-231	Sequence 231, App
7	86	53.8	34	9	US-09-848-834A-13	Sequence 13, Appl
8	79	49.4	34	9	US-09-848-834A-10	Sequence 10, Appl
9	79	49.4	50	9	US-09-848-834A-18	Sequence 18, Appl
10	78	48.8	40	14	US-10-223-711-10	Sequence 10, Appl
11	78	48.8	75	10	US-09-847-102A-31	Sequence 31, Appl
12	78	48.8	75	14	US-10-285-976-229	Sequence 229, App
13	77	48.1	16	9	US-09-848-834A-8	Sequence 8, Appli
14	77	48.1	25	15	US-10-411-544-32	Sequence 32, Appl
15	73	45.6	28	9	US-09-848-834A-11	Sequence 11, Appl
16	73	45.6	33	9	US-09-848-834A-12	Sequence 12, Appl
17	73	45.6	40	14	US-10-223-711-11	Sequence 11, Appl
18	73	45.6	46	9	US-09-848-834A-19	Sequence 19, Appl
19	73	45.6	51	9	US-09-848-834A-20	Sequence 20, Appl
20	72	45.0	15	10	US-09-747-802-16	Sequence 16, Appl
21	72	45.0	15	10	US-09-747-802-30	Sequence 30, Appl
22	72	45.0	15	10	US-09-865-294-8	Sequence 8, Appli
23	72	45.0	15	10	US-09-865-294-22	Sequence 22, Appl
24	72	45.0	15	14	US-10-261-446-20	Sequence 20, Appl
25	72	45.0	15	15	US-10-411-544-10	Sequence 10, Appl
26	72	45.0	36	14	US-10-351-641-505	Sequence 505, App
27	72	45.0	550	9	US-09-873-233A-18	Sequence 18, Appl
28	72	45.0	550	9	US-09-873-233A-20	Sequence 20, Appl
29	69	43.1	15	10	US-09-747-802-37	Sequence 37, Appl
30	69	43.1	15	10	US-09-865-294-29	Sequence 29, Appl
31	69	43.1	16	14	US-10-223-711-3	Sequence 3, Appli
32	69	43.1	19	10	US-09-747-802-48	Sequence 48, Appl
33	69	43.1	19	10	US-09-865-294-40	Sequence 40, Appl
34	66	41.2	19	10	US-09-747-802-51	Sequence 51, Appl
35	66	41.2	19	10	US-09-865-294-43	Sequence 43, Appl
36	65	40.6	662	10	US-09-951-061A-141	Sequence 141, App
37	64	40.0	15	10	US-09-747-802-35	Sequence 35, Appl
38	64	40.0	15	10	US-09-747-802-38	Sequence 38, Appl
39	64	40.0	15	10	US-09-747-802-40	Sequence 40, Appl
40	64	40.0	15	10	US-09-747-802-42	Sequence 42, Appl
41	64	40.0	15	10	US-09-865-294-27	Sequence 27, Appl
42	64	40.0	15	10	US-09-865-294-30	Sequence 30, Appl
43	64	40.0	15	10	US-09-865-294-32	Sequence 32, Appl
44	64	40.0	15	10	US-09-865-294-34	Sequence 34, Appl
45	64	40.0	19	10	US-09-747-802-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the
; OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to ami
; OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated Lysine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
; OTHER INFORMATION: measles virus fusion protein, F
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(22)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (23)..(31)
; OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
; OTHER INFORMATION: GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (31)..(31)

```
;   OTHER INFORMATION: Amidated glycine or glycinamide
US-09-848-834A-9

  Query Match             99.4%;  Score 159;  DB 9;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||||||||||||||||||||||
Db          1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30


RESULT 2
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
;   OTHER INFORMATION: uence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: Amidated-glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (35)..(38)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (39)..(47)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

  Query Match             99.4%;  Score 159;  DB 9;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||||||||||||||||||||||
Db         17 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 46


RESULT 3
US-10-076-674-9
; Sequence 9, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
;  APPLICANT: Sokoll, Kenneth K.
;  TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
;  FILE REFERENCE: Immunogen Delivery System
;  CURRENT APPLICATION NUMBER: US/10/076,674
;  CURRENT FILING DATE:  2002-04-23
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 45
;   TYPE: PRT
;   ORGANISM: Human
US-10-076-674-9

  Query Match             72.5%;  Score 116;  DB 14;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 1.5e-09;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||| |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 4
US-10-355-161A-9
; Sequence 9, Application US/10355161A
; Publication No. US20040009897A1
; GENERAL INFORMATION:
;  APPLICANT: Sokoll, Kenneth K.
;  TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
;  FILE REFERENCE: Immunogen Delivery System
;  CURRENT APPLICATION NUMBER: US/10/355,161A
;  CURRENT FILING DATE:  2003-01-31
;  PRIOR APPLICATION NUMBER: US 10/076674
;  PRIOR FILING DATE: 2002-02-14
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 45
;   TYPE: PRT
;   ORGANISM: Human
US-10-355-161A-9

  Query Match             72.5%;  Score 116;  DB 15;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 1.5e-09;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||| |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 5
US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
;  APPLICANT: University of California
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Rhee, Chae-Seo
;  APPLICANT:  Lorenzo, Leoni M.
;  APPLICANT:  Malini, Sen
;  TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
;  FILE REFERENCE: 22000-20629.00
;  CURRENT APPLICATION NUMBER: US/09/847,102A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
```

```
;   OTHER INFORMATION: PMMVF-FZD2
US-09-847-102A-33

  Query Match             61.9%;  Score 99;  DB 10;  Length 75;
  Best Local Similarity   95.5%;  Pred. No. 7.7e-07;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVEGPSL 22
              |||| ||||||||||||||||||
Db          2 KLLSLIKGVIVHRLEGVEGPSL 23


RESULT 6
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
;  APPLICANT: Rhee, Chae-Seo
;  APPLICANT:  Malini, Sen
;  APPLICANT:  Wu, Christina
;  APPLICANT:  Leoni, Lorenzo M.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
;  TITLE OF INVENTION:  in Head and Neck Squamous Cell Carcinomas
;  FILE REFERENCE: 02307O-130320US
;  CURRENT APPLICATION NUMBER: US/10/285,976
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US 60/287,995
;  PRIOR FILING DATE: 2001-05-01
;  PRIOR APPLICATION NUMBER: WO PCT/US02/13802
;  PRIOR FILING DATE: 2002-05-01
;  NUMBER OF SEQ ID NOS: 232
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PMMVF-ZD2
;   OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
;   OTHER INFORMATION: frizzled domain
US-10-285-976-231

  Query Match             61.9%;  Score 99;  DB 14;  Length 75;
  Best Local Similarity   95.5%;  Pred. No. 7.7e-07;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVEGPSL 22
              |||| ||||||||||||||||||
Db          2 KLLSLIKGVIVHRLEGVEGPSL 23


RESULT 7
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: he Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
;   OTHER INFORMATION: virus fusion protein, F
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

  Query Match             53.8%;  Score 86;  DB 9;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVE 18
              ||||||||||||||||||
Db         17 KLLSEIKGVIVHRLEGVE 34


RESULT 8
US-09-848-834A-10
; Sequence 10, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of t
;   OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylysin) linked by a spacer to a
;   OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Amidated phenylalanine
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(21)
;   OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
;   OTHER INFORMATION: (Tentoxylysin)
;   NAME/KEY: PEPTIDE
;   LOCATION: (22)..(25)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (26)..(34)
;   OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: Amidated glycine or glycinamide
US-09-848-834A-10

  Query Match             49.4%;  Score 79;  DB 9;  Length 34;
  Best Local Similarity   92.9%;  Pred. No. 0.00025;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 VEGPSLHWSYGLRP 30
              :|||||||||||||
```

```
Db         20 LEGPSLHWSYGLRP 33


RESULT 9
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer t
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match             49.4%;  Score 79;  DB 9;  Length 50;
  Best Local Similarity   92.9%;  Pred. No. 0.00039;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 VEGPSLHWSYGLRP 30
              :||||||||||||
Db         36 LEGPSLHWSYGLRP 49


RESULT 10
US-10-223-711-10
; Sequence 10, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
;  APPLICANT: Bakaletz, Lauren O.
;  APPLICANT:  Kaumaya, Pravin T.P.
;  TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
;  FILE REFERENCE: 18525/04058
;  CURRENT APPLICATION NUMBER: US/10/223,711
;  CURRENT FILING DATE:  2002-08-19
;  PRIOR APPLICATION NUMBER: 09/148,711
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 08/460,502
;  PRIOR FILING DATE: 1995-06-02
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-10-223-711-10

  Query Match             48.8%;  Score 78;  DB 14;  Length 40;
  Best Local Similarity   94.4%;  Pred. No. 0.00042;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVE 18
              |||| ||||||||||||||
Db         23 KLLSLIKGVIVHRLEGVE 40


RESULT 11
US-09-847-102A-31
; Sequence 31, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
;  APPLICANT: University of California
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Rhee, Chae-Seo
;  APPLICANT:  Lorenzo, Leoni M.
;  APPLICANT:  Malini, Sen
;  TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
;  FILE REFERENCE: 22000-20629.00
;  CURRENT APPLICATION NUMBER: US/09/847,102A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: PFZD2-MMVF
US-09-847-102A-31

  Query Match             48.8%;  Score 78;  DB 10;  Length 75;
  Best Local Similarity   94.4%;  Pred. No. 0.00085;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVE 18
              |||| ||||||||||||||
Db         58 KLLSLIKGVIVHRLEGVE 75


RESULT 12
US-10-285-976-229
; Sequence 229, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
;  APPLICANT: Rhee, Chae-Seo
;  APPLICANT:  Malini, Sen
;  APPLICANT:  Wu, Christina
;  APPLICANT:  Leoni, Lorenzo M.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
;  TITLE OF INVENTION:  in Head and Neck Squamous Cell Carcinomas
;  FILE REFERENCE: 02307C-130320US
;  CURRENT APPLICATION NUMBER: US/10/285,976
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US 60/287,995
```

```
;  PRIOR FILING DATE: 2001-05-01
;  PRIOR APPLICATION NUMBER: WO PCT/US02/13802
;  PRIOR FILING DATE: 2002-05-01
;  NUMBER OF SEQ ID NOS: 232
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 229
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PFZD2-MMVF
;   OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
;   OTHER INFORMATION: frizzled domain
US-10-285-976-229

  Query Match             48.8%;  Score 78;  DB 14;  Length 75;
  Best Local Similarity   94.4%;  Pred. No. 0.00085;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KLLSEIKGVIVHRLEGVE 18
              |||| ||||||||||||
Db         58 KLLSLIKGVIVHRLEGVE 75


RESULT 13
US-09-848-834A-8
; Sequence 8, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Measles virus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the measles
;   OTHER INFORMATION: virus fusion protein, F
US-09-848-834A-8

  Query Match             48.1%;  Score 77;  DB 9;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVE 18
              ||||||||||||||||
Db          1 LSEIKGVIVHRLEGVE 16


RESULT 14
US-10-411-544-32
; Sequence 32, Application US/10411544
; Publication No. US20030232758A1
; GENERAL INFORMATION:
;  APPLICANT: St. George-Hyslop, Peter
;  APPLICANT:  McLaurin, JoAnne
;  TITLE OF INVENTION: Immunological Methods and Compositions for the Treatment of Alzhe
;  TITLE OF INVENTION:  Disease
;  FILE REFERENCE: LI01547
;  CURRENT APPLICATION NUMBER: US/10/411,544
;  CURRENT FILING DATE:  2003-04-10
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: chimeric sequence
US-10-411-544-32

  Query Match             48.1%;  Score 77;  DB 15;  Length 25;
  Best Local Similarity   72.0%;  Pred. No. 0.00035;
  Matches   18;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYG 27
              |||||||||||||||  |      ||
Db          1 LSEIKGVIVHRLEGVGGFRHDSGYG 25


RESULT 15
US-09-848-834A-11
; Sequence 11, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 28
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of t
;   OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylysin) linked by a spacer to a
;   OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Amidated-glutamine
;   NAME/KEY: MOD_RES
;   LOCATION: (28)..(28)
;   OTHER INFORMATION: Amidated-glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
;   OTHER INFORMATION: (Tentoxylysin)
;   NAME/KEY: PEPTIDE
;   LOCATION: (16)..(19)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (20)..(28)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-11

  Query Match             45.6%;  Score 73;  DB 9;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 GPSLHWSYGLRP 30
              ||||||||||||
Db         16 GPSLHWSYGLRP 27


Search completed: March 10, 2004, 10:25:48
Job time : 25.3658 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 9.64981 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	45.0	282	2	PQ0376	cell fusion glycop
2	72	45.0	282	2	PQ0388	cell fusion glycop
3	72	45.0	534	1	JU0274	cell fusion glycop
4	72	45.0	546	2	S47300	gene F protein - r
5	72	45.0	550	1	E48556	cell fusion glycop
6	72	45.0	553	1	VGNZMV	cell fusion glycop
7	71	44.4	546	1	VGNZRK	cell fusion glycop
8	71	44.4	546	2	S47305	gene F protein - r
9	66.5	41.6	552	2	S47034	cell fusion protei
10	66	41.2	546	1	VGNZRL	cell fusion glycop
11	65	40.6	542	2	JQ2223	cell fusion protei
12	65	40.6	662	1	VGNZCD	cell fusion glycop
13	65	40.6	662	2	S21382	cell fusion protei
14	64	40.0	631	1	VGNZPD	cell fusion glycop
15	64	40.0	631	1	A48346	cell fusion glycop
16	60	37.5	546	2	S55386	cell fusion protei
17	57.5	35.9	92	1	RHHUG	gonadoliberin prec
18	57	35.6	90	1	RHMSG	gonadoliberin prec
19	57	35.6	92	1	RHRTG	gonadoliberin prec
20	56	35.0	636	2	S47299	gene F protein - r
21	54	33.8	67	2	I78541	gonadoliberin prec
22	53	33.1	508	1	VGVNFR	spike glycoprotein
23	52	32.5	10	1	RHPGG	gonadoliberin - pi
24	52	32.5	10	1	RHSHG	gonadoliberin - sh
25	52	32.5	89	2	I51423	gonadoliberin prec
26	52	32.5	379	1	DCBSPK	phosphoribosylamin
27	50.5	31.6	190	2	T37168	probable tetR-fami
28	50	31.2	333	2	T23151	hypothetical prote
29	49	30.6	451	2	AH0063	conserved hypothet
30	49	30.6	582	2	G71500	aspartate-tRNA lig
31	48.5	30.3	393	2	D83589	glutaryl-CoA dehyd
32	48.5	30.3	413	2	B70907	hypothetical prote
33	48	30.0	10	1	RHAQ1	gonadoliberin I -
34	48	30.0	92	2	I50644	gonadoliberin I pr
35	48	30.0	98	2	I50739	gonadotropin-relea
36	48	30.0	502	2	T36589	probable transmemb
37	48	30.0	576	2	AF2361	flavoprotein [impo
38	48	30.0	856	2	I58411	protein-tyrosine k
39	47.5	29.7	222	1	E71024	probable transport
40	47.5	29.7	229	2	JC7219	nuclear protein SR
41	47.5	29.7	586	2	T08293	hypothetical prote
42	47	29.4	339	2	B97755	hypothetical prote
43	47	29.4	388	2	C72710	probable fmu prote
44	47	29.4	452	2	T10882	coproporphyrinogen
45	47	29.4	516	2	T00514	cytochrome P450 ho

ALIGNMENTS

RESULT 1
PQ0376
cell fusion glycoprotein - measles virus (strain TT) (fragment)
C;Species: measles virus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C;Accession: PQ0376
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0376
A;Molecule type: genomic RNA
A;Residues: 1-282 <SCH>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
||||||||||||||||
Db 20 LSEIKGVIVHRLEGV 34

RESULT 2
PQ0388
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)
C;Species: measles virus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C;Accession: PQ0388
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0388
A;Molecule type: genomic RNA
A;Residues: 1-282 <SCH>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
||||||||||||||||
Db 20 LSEIKGVIVHRLEGV 34

RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: subacute sclerosing panencephalitis virus, SSPEV
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C;Accession: JU0274
R;Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A;Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A;Reference number: JU0274; MUID:90385702; PMID:1698327
A;Accession: JU0274
A;Molecule type: mRNA
A;Residues: 1-534 <KOM>
A;Cross-references: EMBL:D10548; NID:g222256; PIDN:BAA01405.1; PID:g222257
A;Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;498-514/Domain: transmembrane #status predicted <TMN>
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
||||||||||||||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
S47300
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C;Accession: S47300; PQ0865
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A;Reference number: S47299
A;Accession: S47300
A;Molecule type: DNA
A;Residues: 1-546 <EVA>
A;Cross-references: EMBL:Z31656; NID:g535406; PIDN:CAA83482.1; PID:g535407
R;Chamberlain, R.W.; Wamwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J
J. Gen. Virol. 74, 2775-2780, 1993
A;Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A;Reference number: PQ0865; MUID:94103786; PMID:8277286
A;Accession: PQ0865
A;Molecule type: mRNA
A;Residues: 86-191 <CHA>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 45.0%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
||||||||||||||||
Db 284 LSEIKGVIVHRLEGV 298

RESULT 5
E48556
cell fusion glycoprotein precursor - measles virus (strain AIK-C)
C;Species: measles virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C;Accession: E48556
R;Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A;Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C
A;Reference number: A48556; MUID:93227570; PMID:8470368
A;Accession: E48556
A;Molecule type: genomic RNA
A;Residues: 1-550 <MOR>
A;Cross-references: GB:S58435; NID:g299460; PIDN:AAB26145.1; PID:g299465
A;Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;113-138/Region: hydrophobic
F;495-514/Domain: transmembrane #status predicted <TMM>
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
||||||||||||||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 6
VGNZMV
cell fusion glycoprotein precursor - measles virus
C;Species: measles virus
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C;Accession: A26962; A25616; PQ0380; PQ0384
R;Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A;Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and comp
A;Reference number: A92794; MUID:87224816; PMID:3585281
A;Accession: A26962
A;Molecule type: mRNA
A;Residues: 1-553 <BUC>
A;Cross-references: GB:D00090; NID:g222061; PIDN:BAA00056.1; PID:g222062
A;Experimental source: strain Halle
R;Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini,
Virology 155, 508-523, 1986
A;Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles vir
A;Reference number: A94350; MUID:87071668; PMID:3788062
A;Accession: A25616
A;Molecule type: mRNA
A;Residues: 4-553 <RIC>
A;Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763
A;Experimental source: strain Edmonston
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparisc
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0380
A;Molecule type: genomic RNA
A;Residues: 272-553 <SCH1>
A;Experimental source: isolate CL
A;Accession: PQ0384
A;Molecule type: genomic RNA
A;Residues: 272-553 <SCH2>
A;Experimental source: isolate SE
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein

```
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;501-517/Domain: transmembrane #status predicted <TMN>
F;32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             45.0%;  Score 72;  DB 1;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 0.033;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 7
VGNZRK
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
C;Accession: A31051
R;Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.
Virology 166, 149-153, 1988
A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis w
A;Reference number: A31051; MUID:88322864; PMID:3413983
A;Accession: A31051
A;Molecule type: genomic RNA
A;Residues: 1-546 <HSU>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>
F;109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>
F;109-134/Domain: transmembrane #status predicted <TN1>
F;491-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             44.4%;  Score 71;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.046;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 8
S47305
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
C;Accession: S47305; S47301
R;Baron, M.D.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The sequence of the N and L genes of Rinderpest virus, and the 5O and 3O
A;Reference number: S47283
A;Accession: S47305
A;Molecule type: mRNA
A;Residues: 1-546 <BAR>
A;Cross-references: EMBL:Z30697; NID:g535396; PIDN:CAA83181.1; PID:g535401; EMBL:Z30700;
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: transmembrane protein

  Query Match             44.4%;  Score 71;  DB 2;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.046;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 9
S47034
cell fusion protein precursor - porpoise morbillivirus
N;Alternate names: F protein
C;Species: porpoise morbillivirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S47034
R;Bolt, G.; Gottschalck, E.; Blixenkkrone-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.;
submitted to the EMBL Data Library, July 1994
A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbil
A;Reference number: S47034
A;Accession: S47034
A;Molecule type: mRNA
A;Residues: 1-552 <BOL>
A;Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640
A;Experimental source: isolate Ulster 88
A;Note: the source is designated as Cetacean morbillivirus
C;Superfamily: parainfluenza virus cell fusion protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-552/Product: fusion protein #status predicted <MAT>

  Query Match             41.6%;  Score 66.5;  DB 2;  Length 552;
  Best Local Similarity   61.5%;  Pred. No. 0.2;
  Matches   16;  Conservative    1;  Mismatches    2;  Indels    7;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGL 28
              |||:|||||||| |       || |
Db        290 LSEVKGVIVHRLEAV-------SYNL 308


RESULT 10
VGNZRL
cell fusion glycoprotein precursor - rinderpest virus (strain L)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C;Accession: A28921
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of th
A;Reference number: A28921; MUID:88219541; PMID:3285575
A;Accession: A28921
A;Molecule type: mRNA
A;Residues: 1-546 <TSU>
A;Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F;105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F;109-133/Domain: transmembrane #status predicted <TN1>
F;485-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             41.2%;  Score 66;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.23;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||| |
Db        284 LSEIKGVIVHRLESV 298


RESULT 11
JQ2223
cell fusion protein F0 precursor - phocine distemper virus
N;Contains: F1 and F2 chains
C;Species: phocine distemper virus
```

C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
C;Accession: JQ2223
R;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oervell
J. Gen. Virol. 74, 1989-1994, 1993
A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites ar
e virus entity.
A;Reference number: JQ2223; MUID:93389459; PMID:8376973
A;Accession: JQ2223
A;Molecule type: mRNA
A;Residues: 1-542 <VIS>
A;Cross-references: GB:L07075
A;Note: the authors translated the codon ATC for residue 4 as Leu
C;Comment: This fusion protein F0 is cleaved into F1 and F2 chains.
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-542/Product: fusion protein #status predicted <MAT>
F;16-99/Product: F2 chain #status predicted <F2C>
F;105-542/Product: F1 chain #status predicted <F1C>
F;105-135/Region: hydrophobic
F;486-512/Domain: transmembrane #status predicted <TMM>
F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 40.6%; Score 65; DB 2; Length 542;
  Best Local Similarity 86.7%; Pred. No. 0.32;
  Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
Db 280 LSEVKGVIVHRLEAV 294

RESULT 12
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N;Contains: fusion protein F1; fusion protein F2
C;Species: canine distemper virus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: JS0321
R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A;Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A;Reference number: JS0321; MUID:88129050; PMID:3433924
A;Accession: JS0321
A;Molecule type: mRNA
A;Residues: 1-662 <BAR>
A;Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-135/Domain: signal sequence #status predicted <SIG>
F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F;606-629/Domain: transmembrane #status predicted <MEM>
F;62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 40.6%; Score 65; DB 1; Length 662;
  Best Local Similarity 86.7%; Pred. No. 0.4;
  Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
Db 400 LSEVKGVIVHRLEAV 414

RESULT 13
S21382
cell fusion protein - canine distemper virus
C;Species: canine distemper virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S21382
R;Wild, T.F.; Bernard, A.; Spehner, D.; Villeval, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A;Description: Vaccination of mice against canine distemper virus induced encephalitis
A;Reference number: S21382
A;Accession: S21382
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-662 <WIL>
A;Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854
C;Superfamily: parainfluenza virus cell fusion protein

  Query Match 40.6%; Score 65; DB 2; Length 662;
  Best Local Similarity 86.7%; Pred. No. 0.4;
  Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
Db 400 LSEVKGVIVHRLEAV 414

RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C;Accession: JQ1368
R;Koevamees, J.; Blixenkrone-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A;Title: The nucleotide sequence and deduced amino acid composition of the haemagglutir
A;Reference number: JQ1368; MUID:92113538; PMID:1765768
A;Accession: JQ1368
A;Molecule type: genomic RNA
A;Residues: 1-631 <KOV>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;189-193/Region: cleavage processing #status predicted
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-212/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 40.0%; Score 64; DB 1; Length 631;
  Best Local Similarity 80.0%; Pred. No. 0.52;
  Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17
Db 369 LSEVKGVVVHRLEAV 383

RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C;Accession: A48346
R;Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A;Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced ami
A;Reference number: A48346; MUID:92398437; PMID:1524494
A;Accession: A48346
A;Molecule type: mRNA
A;Residues: 1-631 <CUR>
A;Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)
C;Genetics:
A;Gene: F

C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-219/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 64; DB 1; Length 631;
Best Local Similarity 80.0%; Pred. No. 0.52;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 LSEIKGVIVHRLEGV 17
              |||:|||:||||| |
Db        369 LSEVKGVVVHRLEAV 383
```

Search completed: March 10, 2004, 09:16:45
Job time : 10.7086 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 5.66926 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	45.0	534	1	VGLF_MEASY	P26032 measles vir
2	72	45.0	546	1	VGLF_RINDB	P41360 rinderpest
3	72	45.0	550	1	VGLF_MEASA	P35973 measles vir
4	72	45.0	550	1	VGLF_MEASE	P08300 measles vir
5	71	44.4	546	1	VGLF_RINDR	P41356 rinderpest
6	66	41.2	546	1	VGLF_RINDL	P10864 rinderpest
7	65	40.6	546	1	VGLF_RINDK	P12574 rinderpest
8	65	40.6	662	1	VGLF_CDVO	P12569 canine dist
9	64	40.0	529	1	VGLF_MEASI	P26031 measles vir
10	64	40.0	631	1	VGLF_PHODV	P28886 phocine dis
11	57.5	35.9	91	1	GON1_PIG	P49921 sus scrofa
12	57.5	35.9	92	1	GON1_HUMAN	P01148 homo sapien
13	57	35.6	90	1	GON1_MOUSE	P13562 mus musculu
14	57	35.6	92	1	GON1_RAT	P07490 rattus norv
15	54	33.8	67	1	GON1_MACMU	P55247 macaca mula
16	53	33.1	508	1	VGLG_IHNV	P07923 infectious
17	52.5	32.8	90	1	GON1_RANCA	Q90y63 rana catesb
18	52	32.5	61	1	GON1_SHEEP	Q28588 ovis aries
19	52	32.5	63	1	GON1_MESAU	O09163 mesocricetu
20	52	32.5	89	1	GON1_XENLA	P45656 xenopus lae
21	52	32.5	92	1	GON1_TUPGB	Q95335 tupaia glis
22	52	32.5	379	1	PURK_BACSU	P12045 bacillus su
23	50.5	31.6	92	1	GON1_CAVPO	O54713 cavia porce
24	50.5	31.6	99	1	GON1_DICLA	Q9ia10 dicentrarch
25	49	30.6	169	1	CX41_THUOB	Q9i8u0 thunnus obe
26	49	30.6	582	1	SYD_CHLTR	O84546 chlamydia t
27	48.5	30.3	95	1	GON1_MORSA	O73812 morone saxa
28	48	30.0	10	1	GON1_ALLMI	P37041 alligator m
29	48	30.0	92	1	GON1_CHICK	P37042 gallus gall
30	48	30.0	94	1	GON1_HAPBU	P51918 haplochromi
31	48	30.0	95	1	GON1_PAGMA	P70074 pagrus majo
32	48	30.0	95	1	GON1_SPAAU	P51919 sparus aura
33	48	30.0	124	1	UCN1_HUMAN	P55089 homo sapien
34	48	30.0	576	1	DFA1_ANASP	Q8ynw5 anabaena sp
35	47	29.4	122	1	UCN1_MOUSE	P81615 mus musculu
36	47	29.4	452	1	HEMF_RHOSH	P33770 rhodobacter
37	47	29.4	452	1	HEMN_RHOSH	P95651 rhodobacter
38	47	29.4	573	1	DFA1_SYNY3	Q55393 synechocyst
39	47	29.4	579	1	DFA2_ANASP	Q8z0c0 anabaena sp
40	46.5	29.1	207	1	TER2_ECOLI	P04483 escherichia
41	46.5	29.1	417	1	RF1_PYRAB	Q9v151 pyrococcus
42	46	28.7	110	1	YHBJ_ACTAC	P96769 actinobacil
43	46	28.7	555	1	VGLF_PI1HC	P12605 human parai
44	46	28.7	652	1	GSH1_ONCVO	Q9nfn6 onchocerca
45	45.5	28.4	338	1	DCUP_AQUAE	O66667 aquifex aeo

ALIGNMENTS

```
RESULT 1
VGLF_MEASY
ID   VGLF_MEASY     STANDARD;      PRT;   534 AA.
AC   P26032;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain Yamagata-1) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90385702; PubMed=1698327;
RA   Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;
RT   "Molecular analysis of structural protein genes of the Yamagata-1
RT   strain of defective subacute sclerosing panencephalitis virus. IV.
RT   Nucleotide sequence of the fusion gene.";
RL   Virus Genes 4:173-181(1990).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10548; BAA01405.1; -.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    534       FUSION GLYCOPROTEIN F0.
FT   CHAIN       24    112       PROTEIN F2.
FT   CHAIN      113    534       PROTEIN F1.
FT   TRANSMEM   113    136       POTENTIAL.
FT   DOMAIN     137    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   495    515       POTENTIAL.
FT   DOMAIN     516    534       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    68    195       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD    29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   534 AA;  57963 MW;  F5B21757E643844D CRC64;
```

```
  Query Match             45.0%;  Score 72;  DB 1;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302
```

```
RESULT 2
VGLF_RINDB
ID   VGLF_RINDB     STANDARD;      PRT;   546 AA.
AC   P41360;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain RBT1) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=39007;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95088609; PubMed=7996154;
RA   Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT   "Nucleotide sequence comparisons of the fusion protein gene from
RT   virulent and attenuated strains of rinderpest virus.";
RL   J. Gen. Virol. 75:3611-3617(1994).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z31656; CAA83482.1; -.
DR   PIR; S47300; S47300.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       O-LINKED (POTENTIAL).
SQ   SEQUENCE   546 AA;  58418 MW;  38B539B89344F401 CRC64;

  Query Match             45.0%;  Score 72;  DB 1;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 0.0062;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
```

```
Db        284 LSEIKGVIVHRLEGV 298
```

```
RESULT 3
VGLF_MEASA
ID   VGLF_MEASA     STANDARD;      PRT;   550 AA.
AC   P35973;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain AIK-C) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=36408;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93227570; PubMed=8470368;
RA   Mori T., Sasaki K., Hashimoto H., Makino S.;
RT   "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT   the AIK-C strain of attenuated measles virus.";
RL   Virus Genes 7:67-81(1993).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S58435; AAB26145.1; -.
DR   PIR; E48556; E48556.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    550       FUSION GLYCOPROTEIN F0.
FT   CHAIN        24    112       PROTEIN F2.
FT   CHAIN       113    550       PROTEIN F1.
FT   TRANSMEM    113    136       POTENTIAL.
FT   DOMAIN      137    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    495    515       POTENTIAL.
FT   DOMAIN      516    550       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     68    195       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  59540 MW;  AAC4DAB92DE0D938 CRC64;

  Query Match             45.0%;  Score 72;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.0062;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302
```

```
RESULT 4
VGLF_MEASE
ID   VGLF_MEASE     STANDARD;      PRT;   550 AA.
AC   P08300;
```

```
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain Edmonston) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Halle) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Leningrad-16) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Edmonston-Zagreb) (Subacute sclerose
OS   panencephalitis virus),
OS   Measles virus (strain Philadelphia-26) (Subacute sclerose
OS   panencephalitis virus), and
OS   Measles virus (strain Edmonston B) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=87071668; PubMed=3788062;
RA   Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
RA   Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT   "The nucleotide sequence of the mRNA encoding the fusion protein of
RT   measles virus (Edmonston strain): a comparison of fusion proteins
RT   from several different paramyxoviruses.";
RL   Virology 155:508-523(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Halle;
RX   MEDLINE=87224816; PubMed=3585281;
RA   Buckland R., Gerald C., Barker R., Wild T.F.;
RT   "Fusion glycoprotein of measles virus: nucleotide sequence of the
RT   gene and comparison with other paramyxoviruses.";
RL   J. Gen. Virol. 68:1695-1703(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=90085790; PubMed=2596022;
RA   Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Baczko K.,
RA   Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT   "Mutated and hypermutated genes of persistent measles viruses which
RT   caused lethal human brain diseases.";
RL   Virology 173:415-425(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=92263801; PubMed=1585658;
RA   Schmid A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,
RA   Billeter M.A.;
RT   "Subacute sclerosing panencephalitis is typically characterized by
RT   alterations in the fusion protein cytoplasmic domain of the
RT   persisting measles virus.";
RL   Virology 188:910-915(1992).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston, Leningrad-16, and Edmonston-Zagreb;
RX   MEDLINE=94249283; PubMed=8191786;
RA   Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT   "Comparison of sequences of the H, F, and N coding genes of measles
RT   virus vaccine strains.";
RL   Virus Res. 31:317-330(1994).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Philadelphia-26;
RX   MEDLINE=94303181; PubMed=8030232;
RA   Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT   "Restriction of fusion protein mRNA as a mechanism of measles virus
RT   persistence.";
RL   Virology 202:665-672(1994).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston B;
RA   Billeter M.A.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14915; AAA46423.1; -.
DR   EMBL; X05597; CAA29090.1; ALT_INIT.
DR   EMBL; K01711; AAA75498.1; ALT_INIT.
DR   EMBL; K01711; AAA75499.1; -.
DR   EMBL; U03657; AAA56647.1; ALT_INIT.
DR   EMBL; U03659; AAA56649.1; ALT_INIT.
DR   EMBL; U03670; AAA56660.1; ALT_INIT.
DR   EMBL; U08416; AAA50550.1; ALT_INIT.
DR   EMBL; Z66517; CAA91367.1; ALT_INIT.
DR   EMBL; Z66517; CAA91368.1; -.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    550       FUSION GLYCOPROTEIN F0.
FT   CHAIN        24    112       PROTEIN F2.
FT   CHAIN       113    550       PROTEIN F1.
FT   TRANSMEM    113    136       POTENTIAL.
FT   DOMAIN      137    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    495    515       POTENTIAL.
FT   DOMAIN      516    550       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     68    195       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  59532 MW;  7AA4F1CA82169093 CRC64;

  Query Match             45.0%;  Score 72;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.0062;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 5
VGLF_RINDR
ID   VGLF_RINDR     STANDARD;      PRT;   546 AA.
AC   P41356;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain RBOK) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=36409;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95088609; PubMed=7996154;
RA   Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT   "Nucleotide sequence comparisons of the fusion protein gene from
RT   virulent and attenuated strains of rinderpest virus.";
RL   J. Gen. Virol. 75:3611-3617(1994).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z30700; CAA83186.1; -.
DR   EMBL; Z30697; CAA83181.1; -.
DR   PIR; S47305; S47305.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   546 AA;  58705 MW;  ED3DF8AFFDBECB95 CRC64;

  Query Match             44.4%;  Score 71;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.0087;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 6
VGLF_RINDL
ID   VGLF_RINDL     STANDARD;      PRT;   546 AA.
AC   P10864;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain L) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11243;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88219541; PubMed=3285575;
RA   Tsukiyama K., Yoshikawa Y., Yamanouchi K.;
RT   "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT   sequence of the F mRNA, and several features of the F protein.";
RL   Virology 164:523-530(1988).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M20870; AAA47399.1; -.
DR   PIR; A28921; VGNZRL.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   546 AA;  58911 MW;  985029418F28FFB5 CRC64;

  Query Match             41.2%;  Score 66;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.048;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||:|
Db        284 LSEIKGVIVHRLESV 298


RESULT 7
VGLF_RINDK
ID   VGLF_RINDK     STANDARD;      PRT;   546 AA.
AC   P12574;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain Kabete O) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11242;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88322864; PubMed=3413983;
RA   Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;
RT   "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT   analysis with other morbilliviruses.";
RL   Virology 166:149-153(1988).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M21514; AAA47400.1; -.
DR   PIR; A31051; VGNZRK.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   546 AA;  58662 MW;  476D74DCC18BCFCF CRC64;

  Query Match             40.6%;  Score 65;  DB 1;  Length 546;
  Best Local Similarity   86.7%;  Pred. No. 0.068;
  Matches   13;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||:| ||||
Db        284 LSEIKGVIIHALEGV 298


RESULT 8
VGLF_CDVO
ID   VGLF_CDVO      STANDARD;      PRT;   662 AA.
AC   P12569; Q65991;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Canine distemper virus (strain Onderstepoort) (CDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11233;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88129050; PubMed=3433924;
RA   Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT   "The nucleotide sequence of the gene encoding the F protein of canine
RT   distemper virus: a comparison of the deduced amino acid sequence with
RT   other paramyxoviruses.";
RL   Virus Res. 8:373-386(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93227696; PubMed=8470428;
RA   Wild T.F., Bernard A., Spehner D., Villeval D., Drillien R.;
RT   "Vaccination of mice against canine distemper virus-induced
RT   encephalitis with vaccinia virus recombinants encoding measles or
RT   canine distemper virus antigens.";
RL   Vaccine 11:438-444(1993).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M21849; AAA42878.1; -.
DR   EMBL; X65509; CAA46481.1; -.
DR   PIR; JS0321; VGNZCD.
DR   PIR; S21382; S21382.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1      ?
FT   CHAIN         ?    662       FUSION GLYCOPROTEIN F0.
FT   CHAIN         ?    224       PROTEIN F2.
FT   CHAIN       225    662       PROTEIN F1.
FT   TRANSMEM    606    629       POTENTIAL.
FT   DISULFID    180    307       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    141    141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    179    179       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      3      3       R -> K (IN REF. 2).
FT   CONFLICT    140    140       D -> N (IN REF. 2).
FT   CONFLICT    152    152       N -> S (IN REF. 2).
FT   CONFLICT    171    171       I -> M (IN REF. 2).
FT   CONFLICT    174    174       A -> V (IN REF. 2).
FT   CONFLICT    662    662       L -> H (IN REF. 2).
SQ   SEQUENCE   662 AA;  72970 MW;  FB2C81C9797805F0 CRC64;

  Query Match             40.6%;  Score 65;  DB 1;  Length 662;
  Best Local Similarity   86.7%;  Pred. No. 0.083;
  Matches   13;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||:|||||||||| |
Db        400 LSEVKGVIVHRLEAV 414


RESULT 9
VGLF_MEASI
ID   VGLF_MEASI     STANDARD;      PRT;   529 AA.
AC   P26031; Q83298;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain IP-3-Ca) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11237;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92263801; PubMed=1585658;
RA   Schmid A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,
RA   Billeter M.A.;
RT   "Subacute sclerosing panencephalitis is typically characterized by
RT   alterations in the fusion protein cytoplasmic domain of the
RT   persisting measles virus.";
RL   Virology 188:910-915(1992).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
```

```
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16566; CAA34567.1; -.
DR   EMBL; X16566; CAA34568.1; ALT_INIT.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    529       FUSION GLYCOPROTEIN F0.
FT   CHAIN        27    115       PROTEIN F2.
FT   CHAIN       116    529       PROTEIN F1.
FT   TRANSMEM    116    139       POTENTIAL.
FT   DOMAIN      140    497       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    498    518       POTENTIAL.
FT   DOMAIN      519    529       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     71    198       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     32     32       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     70     70       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   529 AA;  57331 MW;  AE987BC9F07E9AA9 CRC64;

  Query Match             40.0%;  Score 64;  DB 1;  Length 529;
  Best Local Similarity   93.3%;  Pred. No. 0.093;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||| |||||||||
Db        291 LSEIKEVIVHRLEGV 305


RESULT 10
VGLF_PHODV
ID   VGLF_PHODV     STANDARD;      PRT;   631 AA.
AC   P28886;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Phocine distemper virus (PDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate DK88-4A;
RX   MEDLINE=92113538; PubMed=1765768;
RA   Koevamees J., Blixenkrone-Moeller M., Sharma B., Oervell C.,
RA   Norrby E.;
RT   "The nucleotide sequence and deduced amino acid composition of the
RT   haemagglutinin and fusion proteins of the morbillivirus phocid
RT   distemper virus.";
RL   J. Gen. Virol. 72:2959-2966(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ulster/88;
RX   MEDLINE=92398437; PubMed=1524494;
RA   Curran M.D., Lu Y.J., Rima B.K.;
RT   "The fusion protein gene of phocine distemper virus: nucleotide and
RT   deduced amino acid sequences and a comparison of morbillivirus fusion
RT   proteins.";
RL   Arch. Virol. 126:159-169(1992).
RN   [3]
RP   SEQUENCE OF 95-631 FROM N.A.
RC   STRAIN=Ulster/88;
RX   MEDLINE=91089508; PubMed=2264246;
RA   Curran M.D., Loan D.O., Rima B.K., Kennedy S.;
RT   "Nucleotide sequence analysis of phocine distemper virus reveals its
RT   distinctness from canine distemper virus.";
RL   Vet. Rec. 127:430-431(1990).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D10371; BAA01206.1; -.
DR   PIR; A48346; A48346.
DR   PIR; JQ1368; VGNZPD.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1      ?
FT   CHAIN         ?    631       FUSION GLYCOPROTEIN F0.
FT   CHAIN         ?    188       F2 PROTEIN.
FT   CHAIN       194    631       F1 PROTEIN.
FT   DISULFID    149    276       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   TRANSMEM     89    106       POTENTIAL.
FT   TRANSMEM    194    212       POTENTIAL.
FT   TRANSMEM    575    595       POTENTIAL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     63     63       I -> V (IN REF. 2).
SQ   SEQUENCE   631 AA;  68873 MW;  D1FC87CDD426E9B8 CRC64;

  Query Match             40.0%;  Score 64;  DB 1;  Length 631;
  Best Local Similarity   80.0%;  Pred. No. 0.11;
  Matches   12;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||:|||:||||| |
Db        369 LSEVKGVVVHRLEAV 383


RESULT 11
GON1_PIG
ID   GON1_PIG       STANDARD;      PRT;    91 AA.
AC   P49921;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN   GNRH1 OR GNRH.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RA   Weesner G.D., Matteri R.L., Becker B.A.;
RL   Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
```

```
RP   SEQUENCE OF 24-33.
RX   MEDLINE=72114303; PubMed=4946067;
RA   Baba Y., Matsuo H., Schally A.V.;
RT   "Structure of the porcine LH- and FSH-releasing hormone. II.
RT   Confirmation of the proposed structure by conventional sequential
RT   analyses.";
RL   Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN   [3]
RP   SYNTHESIS OF GONADOLIBERIN.
RX   MEDLINE=72065376; PubMed=4942726;
RA   Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT   "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT   phase method.";
RL   Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN   [4]
RP   SYNTHESIS OF GONADOLIBERIN.
RX   MEDLINE=72117544; PubMed=4946275;
RA   Baba Y., Arimura A., Schally A.V.;
RT   "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
RL   Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L32864; AAA31066.1; -.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23
FT   CHAIN         24    91       PROGONADOLIBERIN I.
FT   PEPTIDE       24    33       GONADOLIBERIN I.
FT   PEPTIDE       34    91       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE      26    26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES       24    24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       33    33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   91 AA;  10090 MW;  8340474F32DDAA99 CRC64;

  Query Match             35.9%;  Score 57.5;  DB 1;  Length 91;
  Best Local Similarity   50.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    4;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||:   |::: |  |   | ||||||||
Db          6 KLLA---GLLLLTLCVVGCSSQHWSYGLRP 32


RESULT 12
GON1_HUMAN
ID   GON1_HUMAN     STANDARD;      PRT;    92 AA.
AC   P01148;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE   peptide I].
GN   GNRH1 OR GNRH OR LHRH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89366682; PubMed=2671939;
RA   Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT   "The complete nucleotide sequence of the human gonadotropin-releasing
RT   hormone gene.";
RL   Nucleic Acids Res. 17:6403-6403(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86094338; PubMed=2867548;
RA   Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT   "Isolation of the gene and hypothalamic cDNA for the common precursor
RT   of gonadotropin-releasing hormone and prolactin release-inhibiting
RT   factor in human and rat.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT SER-16.
RX   MEDLINE=85012739; PubMed=6090951;
RA   Seeburg P.H., Adelman J.P.;
RT   "Characterization of cDNA for precursor of human luteinizing hormone
RT   releasing hormone.";
RL   Nature 311:666-668(1984).
RN   [4]
RP   SEQUENCE OF 24-33.
RX   MEDLINE=83126573; PubMed=6760865;
RA   Tan L., Rousseau P.;
RT   "The chemical identity of the immunoreactive LHRH-like peptide
RT   biosynthesized in the human placenta.";
RL   Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN   [5]
RP   VARIANT SER-16.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [6]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC       Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
CC       (Serono).
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01059; CAA25526.1; -.
DR   EMBL; M12578; AAA35916.1; -.
DR   EMBL; X15215; CAA33285.1; -.
DR   PIR; S05308; RHHUG.
DR   Genew; HGNC:4419; GNRH1.
DR   MIM; 152760; -.
```

```
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Pharmaceutical; Signal; Polymorphism;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT   VARIANT      16     16       W -> S (in dbSNP:6185).
FT                                /FTId=VAR_013943.
SQ   SEQUENCE   92 AA;  10380 MW;  30A72221B076FA79 CRC64;

  Query Match             35.9%;  Score 57.5;  DB 1;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 0.14;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         17 VEG-PSLHWSYGLRP 30
              |||  | ||||||||
Db         18 VEGCSSQHWSYGLRP 32


RESULT 13
GON1_MOUSE
ID   GON1_MOUSE     STANDARD;      PRT;    90 AA.
AC   P13562;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE   I].
GN   GNRH1 OR GNRH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87069928; PubMed=3024317;
RA   Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA   Phillips H.S., Nikolics K., Seeburg P.H.;
RT   "A deletion truncating the gonadotropin-releasing hormone gene is
RT   responsible for hypogonadism in the hpg mouse.";
RL   Science 234:1366-1371(1986).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14872; AAA37717.1; -.
DR   PIR; A47578; RHMSG.
DR   MGD; MGI:95789; Gnrh.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     21
FT   CHAIN        22     90       PROGONADOLIBERIN I.
FT   PEPTIDE      22     31       GONADOLIBERIN I.
FT   PEPTIDE      35     90       PROLACTIN RELEASE-INHIBITING FACTOR I.
FT   ACT_SITE     24     24       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      22     22       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      31     31       AMIDATION (G-32 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   90 AA;  10337 MW;  1C0766FA4826E4D9 CRC64;

  Query Match             35.6%;  Score 57;  DB 1;  Length 90;
  Best Local Similarity   70.6%;  Pred. No. 0.17;
  Matches   12;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    1;

Qy         14 LEGVEGPSLHWSYGLRP 30
              |||    | ||||||||
Db         16 LEGCS--SQHWSYGLRP 30


RESULT 14
GON1_RAT
ID   GON1_RAT       STANDARD;      PRT;    92 AA.
AC   P07490;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE   I].
GN   GNRH1 OR GNRH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86094338; PubMed=2867548;
RA   Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT   "Isolation of the gene and hypothalamic cDNA for the common precursor
RT   of gonadotropin-releasing hormone and prolactin release-inhibiting
RT   factor in human and rat.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89384661; PubMed=2476669;
RA   Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT   "The rat gonadotropin-releasing hormone: SH locus: structure and
RT   hypothalamic expression.";
RL   Mol. Endocrinol. 3:1257-1262(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=93105480; PubMed=1468115;
RA   Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT   "Thymocytes express a mRNA that is identical to hypothalamic
RT   luteinizing hormone-releasing hormone mRNA.";
RL   Cell. Mol. Neurobiol. 12:447-454(1992).
RN   [4]
RP   SEQUENCE OF 1-47 FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=87149087; PubMed=3547652;
RA   Adelman J.P., Bond C.T., Douglass J., Herbert E.;
```

```
RT   "Two mammalian genes transcribed from opposite strands of the same
RT   DNA locus.";
RL   Science 235:1514-1517(1987).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Central nervous system.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S50870; AAB24572.1; -.
DR   EMBL; M12579; AAA41263.1; -.
DR   EMBL; M31670; AAA41264.1; -.
DR   EMBL; M15527; AAA42141.1; ALT_SEQ.
DR   EMBL; M15529; AAA42139.1; -.
DR   EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR   PIR; A40147; RHRTG.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       PROLACTIN RELEASE-INHIBITING FACTOR I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   92 AA;  10500 MW;  494B5C64DA8A3EB3 CRC64;

  Query Match             35.6%;  Score 57;  DB 1;  Length 92;
  Best Local Similarity   70.6%;  Pred. No. 0.17;
  Matches   12;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    1;

Qy         14 LEGVEGPSLHWSYGLRP 30
              |||    | ||||||||
Db         18 LEGCS--SQHWSYGLRP 32


RESULT 15
GON1_MACMU
ID   GON1_MACMU     STANDARD;      PRT;    67 AA.
AC   P55247;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE   (Fragment).
GN   GNRH1 OR GNRH OR LHRH.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RX   MEDLINE=95124501; PubMed=7545971;
RA   Ma Y.J., Costa M.E., Ojeda S.R.;
RT   "Developmental expression of the genes encoding transforming growth
RT   factor alpha and its receptor in the hypothalamus of female rhesus
RT   macaques.";
RL   Neuroendocrinology 60:346-359(1994).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S75918; AAB33096.1; -.
DR   PIR; I78541; I78541.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Signal; Pyrrolidone carboxylic acid.
FT   NON_TER       1      1
FT   SIGNAL       <1      5       BY SIMILARITY.
FT   CHAIN         6    >67       PROGONADOLIBERIN I.
FT   PEPTIDE       6     15       GONADOLIBERIN I.
FT   PEPTIDE      19    >67       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE      8      8       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY (BY SIMILARITY).
FT   MOD_RES       6      6       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   MOD_RES      15     15       AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT                                SIMILARITY).
FT   NON_TER      67     67
SQ   SEQUENCE   67 AA;  7573 MW;  505394DAA261A3F2 CRC64;

  Query Match             33.8%;  Score 54;  DB 1;  Length 67;
  Best Local Similarity   90.0%;  Pred. No. 0.35;
  Matches    9;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         21 SLHWSYGLRP 30
              | ||||||||
Db          5 SQHWSYGLRP 14


Search completed: March 10, 2004, 09:13:52
Job time : 5.66926 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 30.6381 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	45.0	534	12	Q04243	Q04243 measles vir
2	72	45.0	537	12	Q04242	Q04242 measles vir
3	72	45.0	545	12	Q9PXA4	Q9pxa4 measles vir
4	72	45.0	546	12	Q91HA5	Q91ha5 rinderpest
5	72	45.0	550	12	P90331	P90331 measles vir
6	72	45.0	550	12	Q9QEX0	Q9qex0 measles vir
7	72	45.0	550	12	Q9QEW9	Q9qew9 measles vir
8	72	45.0	550	12	P90330	P90330 measles vir
9	72	45.0	550	12	Q9QEW7	Q9qew7 measles vir
10	72	45.0	550	12	Q9WMK4	Q9wmk4 measles vir
11	72	45.0	550	12	Q89495	Q89495 measles vir
12	72	45.0	550	12	Q8V049	Q8v049 measles vir
13	72	45.0	550	12	Q9YJ94	Q9yj94 measles vir
14	72	45.0	550	12	Q9QEX1	Q9qex1 measles vir
15	72	45.0	550	12	Q9QEW8	Q9qew8 measles vir
16	72	45.0	553	12	O93055	O93055 measles vir
17	72	45.0	553	12	Q9IC36	Q9ic36 measles vir
18	72	45.0	553	12	P88973	P88973 measles vir
19	72	45.0	553	12	Q83536	Q83536 measles vir
20	72	45.0	553	12	O11383	O11383 measles vir
21	72	45.0	553	12	Q9IFK2	Q9ifk2 measles vir
22	72	45.0	553	12	Q83533	Q83533 measles vir
23	72	45.0	553	12	Q83525	Q83525 measles vir
24	72	45.0	553	12	Q83518	Q83518 measles vir
25	72	45.0	553	12	P88974	P88974 measles vir
26	72	45.0	553	12	Q83527	Q83527 measles vir
27	72	45.0	553	12	Q83521	Q83521 measles vir
28	72	45.0	553	12	Q83530	Q83530 measles vir
29	72	45.0	553	12	O91248	O91248 measles vir
30	72	45.0	553	12	Q91QP2	Q91qp2 measles vir
31	72	45.0	553	12	Q04244	Q04244 measles vir
32	72	45.0	579	12	Q9PWU4	Q9pwu4 measles vir
33	68	42.5	545	12	Q9QEW6	Q9qew6 measles vir
34	68	42.5	553	12	O11380	O11380 measles vir
35	66.5	41.6	552	12	Q66147	Q66147 cetacean mo
36	65	40.6	528	12	Q9YJW9	Q9yjw9 canine dist
37	65	40.6	530	12	Q8QV06	Q8qv06 canine dist
38	65	40.6	662	12	Q9DXZ2	Q9dxz2 canine dist
39	65	40.6	662	12	Q91KN3	Q91kn3 canine dist
40	65	40.6	662	12	Q9YKL7	Q9ykl7 canine dist
41	65	40.6	662	12	O89327	O89327 canine dist
42	63.5	39.7	552	12	Q66409	Q66409 dolphin mor
43	63.5	39.7	552	12	O56852	O56852 dolphin mor
44	61	38.1	553	12	Q83629	Q83629 measles vir
45	60	37.5	546	12	Q84926	Q84926 peste-des-p

ALIGNMENTS

RESULT 1
Q04243
ID Q04243 PRELIMINARY; PRT; 534 AA.
AC Q04243;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16568; CAA34581.1; -.
DR EMBL; X16568; CAA34582.1; -.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 534 AA; 57899 MW; 637245E23B5BE044 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17

```
                  ||||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 2
Q04242
ID   Q04242      PRELIMINARY;      PRT;   537 AA.
AC   Q04242;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89003063; PubMed=3167982;
RA   Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA   Billeter M.A.;
RT   "Biased hypermutation and other genetic changes in defective measles
RT   viruses in human brain infections.";
RL   Cell 55:255-265(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Cattaneo R., Billeter M.A.;
RL   Virology 0:0-0(0).
DR   EMBL; X16567; CAA34574.1; -.
DR   EMBL; X16567; CAA34575.1; -.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   537 AA;  58275 MW;  D0A60AC66D979E06 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 537;
  Best Local Similarity   100.0%;  Pred. No. 0.073;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 3
Q9PXA4
ID   Q9PXA4      PRELIMINARY;      PRT;   545 AA.
AC   Q9PXA4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-3;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179440; AAF02705.1; -.
DR   EMBL; AF179439; AAF02704.1; -.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   545 AA;  58907 MW;  0234C28AE193E77D CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 4
Q91HA5
ID   Q91HA5      PRELIMINARY;      PRT;   546 AA.
AC   Q91HA5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Rinderpest virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K;
RX   MEDLINE=21014265; PubMed=11186456;
RA   Aianot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA   Gusev A.A.;
RT   "Primary structure of the F-gene from Rinderpest virus strain K.";
RL   Mol. Gen. Mikrobiol. Virusol. 4:29-33(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K;
RA   Ayanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA   Gusev A.A.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY035887; AAK63190.1; -.
DR   PIR; PQ0866; PQ0866.
DR   PIR; PQ0867; PQ0867.
DR   PIR; PQ0873; PQ0873.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   546 AA;  58572 MW;  449B2B2DD7405F0B CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 0.074;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              ||||||||||||||||
Db        284 LSEIKGVIVHRLEGV 298


RESULT 5
P90331
ID   P90331      PRELIMINARY;      PRT;   550 AA.
AC   P90331;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Sheng J., Watanabe M., Ueda S.;
RT   "Selection of a neurotropic variant of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT   "An amino acid alteration of F protein responsible for the enhanced
RT   fusogenicity of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D63926; BAA09958.1; -.
DR   EMBL; AF179431; AAF02696.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59530 MW;  97C991C7E2169839 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 6
Q9QEX0
ID   Q9QEX0      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEX0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Toyoshima;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179432; AAF02697.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59504 MW;  2AA969D37FA5CA17 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 7
Q9QEW9
ID   Q9QEW9      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179436; AAF02701.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59405 MW;  0AE6DBFC5DD22BBA CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 8
P90330
ID   P90330      PRELIMINARY;      PRT;   550 AA.
AC   P90330;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nagahata(HB);
RA   Sheng J., Watanabe M., Ueda S.;
RT   "Selection of a neurotropic variant of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nagahata(HB);
RA   Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT   "An amino acid alteration of F protein responsible for the enhanced
```

```
RT   fusogenicity of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D63924; BAA09951.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59589 MW;  73E7BD457ABA39B7 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 9
Q9QEW7
ID   Q9QEW7      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179438; AAF02703.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59333 MW;  086E51FED5582BBA CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 10
Q9WMK4
ID   Q9WMK4      PRELIMINARY;      PRT;   550 AA.
AC   Q9WMK4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WTF;
RX   MEDLINE=99329215; PubMed=10400788;
RA   Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
RA   Schneider-Schaulies S.;
RT   "A recombinant measles vaccine virus expressing wild-type
RT   glycoproteins : consequences for viral spread and cell tropism.";
RL   J. Virol. 73:6903-6915(1999).
DR   EMBL; AJ133108; CAB38075.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59580 MW;  8255499968B5D862 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 11
Q89495
ID   Q89495      PRELIMINARY;      PRT;   550 AA.
AC   Q89495;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92230209; PubMed=1566568;
RA   Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT   "Genetic variability of the glycoprotein genes of current wild-type
RT   measles isolates.";
RL   Virology 188:135-142(1992).
DR   EMBL; M81903; AAA46422.1; -.
DR   EMBL; M81901; AAA46421.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59564 MW;  A78EEC9CD6268E58 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 12
Q8V049
ID   Q8V049      PRELIMINARY;      PRT;   550 AA.
AC   Q8V049;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G954;
RX   MEDLINE=21635526; PubMed=11773423;
RA   Waku Kouomou D., Wild T.F.;
RT   "Adaptation of wild-type measles virus to tissue culture.";
RL   J. Virol. 76:1505-1509(2002).
DR   EMBL; AY059392; AAL29688.1; -.
DR   PIR; PQ0376; PQ0376.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59551 MW;  9A7A4BA99E4DA8E9 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 13
Q9YJ94
ID   Q9YJ94      PRELIMINARY;      PRT;   550 AA.
AC   Q9YJ94;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9301V;
RX   MEDLINE=98440529; PubMed=9765410;
RA   Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA   Asakawa M., Nagai Y.;
RT   "Measles virus attenuation associated with transcriptional impediment
RT   and a few amino acid changes in the polymerase and accessory
RT   proteins.";
RL   J. Virol. 72:8690-8696(1998).
DR   EMBL; AB012949; BAA33877.1; -.
DR   EMBL; AB012948; BAA33871.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59512 MW;  7AA4F1D117197BF9 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 14
Q9QEX1
ID   Q9QEX1      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEX1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Masusako;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179430; AAF02695.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59559 MW;  609EE024A7E59C54 CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 15
Q9QEW8
ID   Q9QEW8      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179437; AAF02702.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59315 MW;  086E51FED235EBBA CRC64;

  Query Match             45.0%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          3 LSEIKGVIVHRLEGV 17
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302
```

Search completed: March 10, 2004, 09:25:31
Job time : 32.6381 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 46.6809 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	159	99.4	31	5	AAU11420	Aau11420 Synthetic
2	159	99.4	47	5	AAU11428	Aau11428 Synthetic
3	117	73.1	25	2	AAR62705	Aar62705 LHRH-cont
4	117	73.1	42	2	AAR62708	Aar62708 LHRH-cont
5	116	72.5	27	2	AAR62707	Aar62707 LHRH-cont
6	116	72.5	27	3	AAY68567	Aay68567 Peptide i
7	116	72.5	27	3	AAY91156	Aay91156 MVF Th ep
8	116	72.5	45	2	AAR62721	Aar62721 LHRH-cont
9	116	72.5	45	7	ADD89949	Add89949 LHRH pept
10	113	70.6	27	3	AAY91163	Aay91163 Modified
11	110	68.8	31	3	AAY91175	Aay91175 Modified
12	108	67.5	27	3	AAY91161	Aay91161 Modified
13	108	67.5	27	3	AAY91167	Aay91167 Modified
14	108	67.5	45	3	AAY68573	Aay68573 Peptide i
15	108	67.5	45	3	AAY91165	Aay91165 Modified
16	107	66.9	31	3	AAY91179	Aay91179 Modified
17	106	66.2	28	3	AAY91158	Aay91158 Modified
18	105	65.6	31	3	AAY68582	Aay68582 Peptide i
19	105	65.6	31	3	AAY91173	Aay91173 Modified
20	105	65.6	47	3	AAY68583	Aay68583 Peptide i
21	105	65.6	47	3	AAY91180	Aay91180 Inv epito
22	105	65.6	49	3	AAY91177	Aay91177 Modified
23	103	64.4	27	3	AAY68575	Aay68575 Peptide i
24	103	64.4	27	3	AAY91170	Aay91170 Modified
25	103	64.4	35	3	AAY91242	Aay91242 Modified
26	100	62.5	28	2	AAR62726	Aar62726 LHRH-cont
27	100	62.5	46	2	AAR62728	Aar62728 LHRH-cont
28	100	62.5	47	3	AAY68586	Aay68586 Peptide i
29	100	62.5	47	3	AAY91183	Aay91183 Inv epito
30	99	61.9	28	3	AAY91159	Aay91159 Modified
31	99	61.9	75	6	ABP72236	Abp72236 Frizzled
32	94	58.8	28	3	AAY91157	Aay91157 Modified
33	89	55.6	27	2	AAR74261	Aar74261 SSAL1 TH1
34	87	54.4	34	2	AAW05619	Aaw05619 MVF1Th-GG
35	86	53.8	34	5	AAU11424	Aau11424 Synthetic
36	84	52.5	27	3	AAY68576	Aay68576 Peptide i
37	84	52.5	27	3	AAY91171	Aay91171 Modified
38	84	52.5	45	3	AAY68577	Aay68577 Peptide i
39	84	52.5	45	3	AAY91172	Aay91172 Modified
40	81	50.6	31	3	AAY68587	Aay68587 Peptide i
41	81	50.6	31	3	AAY91184	Aay91184 Modified
42	79	49.4	27	3	AAY68574	Aay68574 Peptide i
43	79	49.4	27	3	AAY91168	Aay91168 Modified
44	79	49.4	27	3	AAY91164	Aay91164 Modified
45	79	49.4	27	3	AAY91169	Aay91169 Modified

ALIGNMENTS

RESULT 1
AAU11420
ID AAU11420 standard; peptide; 31 AA.
XX
AC AAU11420;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 1.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Measles virus.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .18
FT /note= "Measles virus fusion protein F epitope"
FT Peptide 19. .22
FT /note= "Spacer peptide"
FT Peptide 23. .31
FT /note= "Gonadotrophin releasing hormone epitope"
FT Modified-site 31
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT or its analog.
XX
PS Claim 11; Page 7; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 31 AA;

Query Match 99.4%; Score 159; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.3e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
Db 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30

RESULT 2
AAU11428
ID AAU11428 standard; peptide; 47 AA.
XX
AC AAU11428;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 9.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope (1. .10
FT aa)"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11. .16
FT /note= "Spacer peptide"
FT Peptide 17. .34
FT /note= "Malaria CSP protein (288-302 aa)"
FT Peptide 35. .38
FT /note= "Spacer peptide"
FT Peptide 39. .47
FT /note= "Gonadotrophin releasing hormone epitope (2-10
FT aa)"
FT Modified-site 47
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 11; Page 11; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 47 AA;

Query Match 99.4%; Score 159; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
Db 17 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 46

RESULT 3
AAR62705
ID AAR62705 standard; peptide; 25 AA.
XX
AC AAR62705;
XX
DT 25-MAR-2003 (revised)
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW measles virus F protein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1. .15
FT /note= "measles virus F protein helper T cell epitope"
FT Domain 16. .25
FT /note= "LHRH hapten"
XX
PN WO9425060-A1.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US004832.
XX
PR 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX

```
PA   (LADD/) LADD A E.
PA   (WANG/) WANG C Y.
PA   (ZAMB/) ZAMB T.
XX
PI   Ladd AE,  Wang CY,  Zamb T;
XX
DR   WPI; 1994-357910/44.
XX
PT   Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT   suppress LHRH activity in males and females.
XX
PS   Claim 8; Page 84; 213pp; English.
XX
CC   Synthetic immunogenic peptides are provided in which a universal immune
CC   stimulator is linked to a peptide or protein hapten containing B cell
CC   and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC   potent immune responses to the coupled peptide or protein. The stimulator
CC   consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC   immune response to the coupled peptide in members of a heterogeneous
CC   population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC   sequence from the invasin protein of Yersinia. Spacer amino acid
CC   sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC   domains and between the immune stimulator and hapten components. When the
CC   hapten is LHRH, then optionally the invasin domain can be omitted from
CC   the immune stimulator component. The present sequence represents an LHRH-
CC   containing, invasin-free immunogenic peptide as above which can be used
CC   as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC   dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC   endometriosis, benign uterine tumours, recurrent functional ovarian
CC   cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC   cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 25 AA;

  Query Match             73.1%;  Score 117;  DB 2;  Length 25;
  Best Local Similarity   85.7%;  Pred. No. 1.7e-10;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||||    ||||||||
Db          1 LSEIKGVIVHRLEGVE----HWSYGLRP 24


RESULT 4
AAR62708
ID   AAR62708 standard; peptide; 42 AA.
XX
AC   AAR62708;
XX
DT   25-MAR-2003  (revised)
DT   10-SEP-1995  (first entry)
XX
DE   LHRH-containing immunogenic peptide.
XX
KW   Helper T cell epitope; universal immune stimulator; invasin; hapten;
KW   vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW   androgen-dependent carcinoma; antitumour; infertility;
KW   measles virus F protein.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .15
FT                   /note= "measles virus F protein helper T cell epitope"
FT   Domain          16. .30
FT                   /note= "measles virus F protein helper T cell epitope"
FT   Domain          33. .42
FT                   /note= "LHRH hapten"
XX
PN   WO9425060-A1.
XX
PD   10-NOV-1994.
XX
PF   28-APR-1994;   94WO-US004832.
XX
PR   27-APR-1993;   93US-00057166.
PR   14-APR-1994;   94US-00229275.
XX
PA   (LADD/) LADD A E.
PA   (WANG/) WANG C Y.
PA   (ZAMB/) ZAMB T.
XX
PI   Ladd AE,  Wang CY,  Zamb T;
XX
DR   WPI; 1994-357910/44.
XX
PT   Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT   suppress LHRH activity in males and females.
XX
PS   Claim 8; Page 86; 213pp; English.
XX
CC   Synthetic immunogenic peptides are provided in which a universal immune
CC   stimulator is linked to a peptide or protein hapten containing B cell
CC   and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC   potent immune responses to the coupled peptide or protein. The stimulator
CC   consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC   immune response to the coupled peptide in members of a heterogeneous
CC   population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC   sequence from the invasin protein of Yersinia. Spacer amino acid
CC   sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC   domains and between the immune stimulator and hapten components. When the
CC   hapten is LHRH, then optionally the invasin domain can be omitted from
CC   the immune stimulator component. The present sequence represents an LHRH-
CC   containing, invasin-free immunogenic peptide as above which can be used
CC   as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC   dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC   endometriosis, benign uterine tumours, recurrent functional ovarian
CC   cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC   cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 42 AA;

  Query Match             73.1%;  Score 117;  DB 2;  Length 42;
  Best Local Similarity   82.8%;  Pred. No. 3.2e-10;
  Matches   24;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          2 LLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||||||| |   ||||||||
Db         15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41


RESULT 5
AAR62707
ID   AAR62707 standard; peptide; 27 AA.
XX
AC   AAR62707;
XX
DT   25-MAR-2003  (revised)
DT   10-SEP-1995  (first entry)
XX
DE   LHRH-containing immunogenic peptide.
XX
KW   Helper T cell epitope; universal immune stimulator; invasin; hapten;
KW   vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW   androgen-dependent carcinoma; antitumour; infertility;
KW   measles virus F protein.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .15
FT                   /note= "measles virus F protein helper T cell epitope"
FT   Domain          18. .27
FT                   /note= "LHRH hapten"
```

XX
PN WO9425060-A1.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US004832.
XX
PR 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX
PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX
PI Ladd AE, Wang CY, Zamb T;
XX
DR WPI; 1994-357910/44.
XX
PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
PS Claim 8, 12; Page 86; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasin protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasin domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing, invasin-free immunogenic peptide as above which can be used
CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC endometriosis, benign uterine tumours, recurrent functional ovarian
CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC cancer, or for induction of infertility. This sequence is particularly
CC preferred. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 27 AA;

Query Match 72.5%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.7e-10;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
||||||||||||||| | |||||||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 6
AAY68567
ID AAY68567 standard; peptide; 27 AA.
XX
AC AAY68567;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX
KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; chimera;
KW immunocastration.
XX
OS Measles virus.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .15
FT /note= "helper Th epitope AAY68540"
FT Peptide 16. .17
FT /note= "spacer"
FT Peptide 18. .27
FT /note= "LHRH antigenic epitope AAY68566"
XX
PN WO9966952-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013960.
XX
PR 20-JUN-1998; 98US-00100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160562/14.
XX
PT New peptide immunogen containing luteinizing hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer.
XX
PS Example 1; Page 63; 102pp; English.
XX
CC The present sequence represents a peptide immunogen comprising a helper T
CC cell (Th) epitope of the F protein of the Measles virus and a target
CC antigen, luteinising hormone-releasing hormone (LHRH). The peptide
CC immunogens cause induction of a specific immune response to LHRH which is
CC involved in regulation of spermatogenesis, ovulation, oestrus, sexual
CC development and secretion of sex hormones. Provision of a promiscuous T
CC helper epitope (which is functional in genetically diverse subjects)
CC provides optimum immunogenicity to the B cell epitopes of the target
CC antigen and thus high antibody titres against the target antigen. The
CC peptide immunogens of the invention are used to vaccinate against
CC mammalian LHRH, for use as (reversible) contraceptive; control of hormone
CC -dependent tumours (cancer of prostate or breast, also endometriosis); to
CC prevent boar taint (and improve meat quality) and for immunocastration
XX
SQ Sequence 27 AA;

Query Match 72.5%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.7e-10;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
||||||||||||||| | |||||||
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 7
AAY91156
ID AAY91156 standard; peptide; 27 AA.
XX
AC AAY91156;
XX
DT 12-SEP-2003 (revised)
DT 22-MAY-2000 (first entry)
XX
DE MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX
OS Measles virus.
OS Rattus sp.
OS Chimeric.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013975.
XX
PR 20-JUN-1998; 98US-00100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus.
XX
PS Example 1; Page 77; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and immunocastration)
CC ; for promoting the growth of animals; or for treating allergies or
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC genetically diverse subjects) into an immunogen improves capacity to
CC induce a strong T helper cell-mediated immune response, resulting in
CC production of antibodies against a target antigen. Th can replace carrier
CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC represents a promiscuous T helper epitope from the measles virus F (MVF)
CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC epitope. Somatostatin immunogens may be used to promote growth in
CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC be used to prevent HIV infection of T cells. AAY90212 is a modified
CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC immunogens comprising a CETP peptide and a Th epitope which may be used
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC both of which may optionally be used in the antigenic peptides of the
CC invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 27 AA;

```
  Query Match             72.5%;  Score 116;  DB 3;  Length 27;
  Best Local Similarity   85.7%;  Pred. No. 2.7e-10;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||| |   |||||||
Db          1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26
```

RESULT 8
AAR62721
ID AAR62721 standard; peptide; 45 AA.
XX
AC AAR62721;
XX
DT 25-MAR-2003 (revised)
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW measles virus F protein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1. .16
FT /note= "invasin domain"
FT Domain 19. .33
FT /note= "measles virus F protein helper T cell epitope"
FT Domain 36. .45
FT /note= "LHRH hapten"
XX
PN WO9425060-A1.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US004832.
XX
PR 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX
PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX
PI Ladd AE, Wang CY, Zamb T;
XX
DR WPI; 1994-357910/44.
XX
PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
PS Claim 8; Page 88; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasin protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasin domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-

```
CC   containing immunogenic peptide as above which can be used as a potent
CC   vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
CC   carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis,
CC   benign uterine tumours, recurrent functional ovarian cysts, (severe)
CC   premenstrual syndrome or oestrogen-dependent breast cancer, or for
CC   induction of infertility. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 45 AA;

  Query Match             72.5%;  Score 116;  DB 2;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 5e-10;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||||| |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 9
ADD89949
ID   ADD89949 standard; protein; 45 AA.
XX
AC   ADD89949;
XX
DT   29-JAN-2004  (first entry)
XX
DE   LHRH peptide used in immunostimulant complex for prostate cancer vaccine.
XX
KW   Immunostimulant; vaccine; human; immunogen; LHRH; immunotherapy;
KW   prostate cancer.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO2003068169-A2.
XX
PD   21-AUG-2003.
XX
PF   14-FEB-2003; 2003WO-US004711.
XX
PR   14-FEB-2002; 2002US-00076674.
PR   31-JAN-2003; 2003US-00076674.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Sokoll KK;
XX
DR   WPI; 2003-778890/73.
XX
PT   Stabilized immunostimulating complex, useful for vaccination, e.g.
PT   against human immune deficiency viruses, comprises cationic peptide
PT   immunogen and anionic oligonucleotide.
XX
PS   Claim 17; SEQ ID NO 9; 159pp; English.
XX
CC   The present sequence is that of a synthetic immunogenic peptide derived
CC   from human LHRH. This is an example of peptides that can be used in
CC   claimed immunostimulatory complexes of the invention that are
CC   specifically adapted to act as adjuvant and as peptide immunogen
CC   stabiliser. The complexes comprise a CpG oligonucleotide and a
CC   biologically active peptide immunogen. The complex is particulate and can
CC   efficiently present peptide immunogens to the cells of the immune system
CC   to produce an immune response. The complexes may be prepared with various
CC   ratios of peptides to CpG oligonucleotides to provide different physical
CC   properties, such as the size of the microparticle. An immunostimulatory
CC   complex comprising the present LHRH derived peptide can be used in a
CC   vaccine for prostate cancer.
XX
SQ   Sequence 45 AA;

  Query Match             72.5%;  Score 116;  DB 7;  Length 45;
  Best Local Similarity   85.7%;  Pred. No. 5e-10;
  Matches   24;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              |||||||||||||||| |   ||||||||
Db         19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44


RESULT 10
AAY91163
ID   AAY91163 standard; peptide; 27 AA.
XX
AC   AAY91163;
XX
DT   12-SEP-2003  (revised)
DT   22-MAY-2000  (first entry)
XX
DE   Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Measles virus.
OS   Rattus sp.
OS   Chimeric.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US013975.
XX
PR   20-JUN-1998;   98US-00100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus.
XX
PS   Example 1; Page 80; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and immunocastration)
CC   ; for promoting the growth of animals; or for treating allergies or
CC   arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC   genetically diverse subjects) into an immunogen improves capacity to
CC   induce a strong T helper cell-mediated immune response, resulting in
CC   production of antibodies against a target antigen. Th can replace carrier
CC   proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC   represents a promiscuous T helper epitope from the measles virus F (MVF)
CC   protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC   represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC   AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC   surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC   derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
```

```
CC   Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC   promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC   used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC   AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC   epitope. Somatostatin immunogens may be used to promote growth in
CC   livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC   AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC   be used to prevent HIV infection of T cells. AAY90212 is a modified
CC   version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC   are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC   treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC   disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC   peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC   circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC   antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC   AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC   immunogens comprising a CETP peptide and a Th epitope which may be used
CC   to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC   and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC   Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC   HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC   vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC   invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC   both of which may optionally be used in the antigenic peptides of the
CC   invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 27 AA;

  Query Match             70.6%;  Score 113;  DB 3;  Length 27;
  Best Local Similarity   82.1%;  Pred. No. 7.7e-10;
  Matches   23;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:|||| |   ||||||||
Db          1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26


RESULT 11
AAY91175
ID   AAY91175 standard; peptide; 31 AA.
XX
AC   AAY91175;
XX
DT   12-SEP-2003  (revised)
DT   22-MAY-2000  (first entry)
XX
DE   Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:55.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Measles virus.
OS   Rattus sp.
OS   Chimeric.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US013975.
XX
PR   20-JUN-1998;   98US-00100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus.
XX
PS   Example 1; Page 84; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and immunocastration)
CC   ; for promoting the growth of animals; or for treating allergies or
CC   arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC   genetically diverse subjects) into an immunogen improves capacity to
CC   induce a strong T helper cell-mediated immune response, resulting in
CC   production of antibodies against a target antigen. Th can replace carrier
CC   proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC   represents a promiscuous T helper epitope from the measles virus F (MVF)
CC   protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC   represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC   AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC   surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC   derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
CC   Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC   promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC   used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC   AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC   epitope. Somatostatin immunogens may be used to promote growth in
CC   livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC   AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC   be used to prevent HIV infection of T cells. AAY90212 is a modified
CC   version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC   are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC   treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC   disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC   peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC   circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC   antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC   AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC   immunogens comprising a CETP peptide and a Th epitope which may be used
CC   to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC   and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC   Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC   HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC   vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC   invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC   both of which may optionally be used in the antigenic peptides of the
CC   invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 31 AA;

  Query Match             68.8%;  Score 110;  DB 3;  Length 31;
  Best Local Similarity   78.6%;  Pred. No. 2.6e-09;
  Matches   22;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              ||||||||||:||||     ||||||||
Db          3 LSEIKGVIVHKLEGVLFGGEHWSYGLRP 30


RESULT 12
AAY91161
ID   AAY91161 standard; peptide; 27 AA.
XX
AC   AAY91161;
```

XX
DT 12-SEP-2003 (revised)
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:41.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Measles virus.
OS Rattus sp.
OS Chimeric.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013975.
XX
PR 20-JUN-1998; 98US-00100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus.
XX
PS Example 1; Page 79; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and immunocastration)
CC ; for promoting the growth of animals; or for treating allergies or
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC genetically diverse subjects) into an immunogen improves capacity to
CC induce a strong T helper cell-mediated immune response, resulting in
CC production of antibodies against a target antigen. Th can replace carrier
CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC represents a promiscuous T helper epitope from the measles virus F (MVF)
CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC epitope. Somatostatin immunogens may be used to promote growth in
CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC be used to prevent HIV infection of T cells. AAY90212 is a modified
CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC immunogens comprising a CETP peptide and a Th epitope which may be used
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC both of which may optionally be used in the antigenic peptides of the
CC invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 27 AA;

Query Match 67.5%; Score 108; DB 3; Length 27;
Best Local Similarity 71.4%; Pred. No. 4.4e-09;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```
Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||::||: |   |||||||
Db          1 ISEIKGVIVHKIEGIGGE--HWSYGLRP 26
```

RESULT 13
AAY91167
ID AAY91167 standard; peptide; 27 AA.
XX
AC AAY91167;
XX
DT 12-SEP-2003 (revised)
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:47.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Measles virus.
OS Rattus sp.
OS Chimeric.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013975.
XX
PR 20-JUN-1998; 98US-00100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus.
XX
PS Example 1; Page 81; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and immunocastration)
CC ; for promoting the growth of animals; or for treating allergies or
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC genetically diverse subjects) into an immunogen improves capacity to
CC induce a strong T helper cell-mediated immune response, resulting in
CC production of antibodies against a target antigen. Th can replace carrier
CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC represents a promiscuous T helper epitope from the measles virus F (MVF)
CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC epitope. Somatostatin immunogens may be used to promote growth in
CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC be used to prevent HIV infection of T cells. AAY90212 is a modified
CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC immunogens comprising a CETP peptide and a Th epitope which may be used
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC both of which may optionally be used in the antigenic peptides of the
CC invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 27 AA;

Query Match 67.5%; Score 108; DB 3; Length 27;
Best Local Similarity 75.0%; Pred. No. 4.4e-09;
Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

```
Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||:|||: |   ||||||||
Db          1 MSEIKGVIVHKLEGMGGE--HWSYGLRP 26
```

RESULT 14
AAY68573
ID AAY68573 standard; peptide; 45 AA.
XX
AC AAY68573;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX
KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; invasin domain;
KW immunocastration.
XX
OS Synthetic.
OS Yersinia sp.
OS Measles virus.
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Peptide 1. .16
FT /note= "invasin domain AAY68565"
FT Peptide 17. .18
FT /note= "spacer"
FT Peptide 19. .33
FT /note= "helper Th epitope AAY68544"
FT Peptide 34. .35
FT /note= "spacer"
FT Peptide 36. .45
FT /note= "LHRH antigenic epitope AAY68566"
XX
PN WO9966952-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013960.
XX
PR 20-JUN-1998; 98US-00100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160562/14.
XX
PT New peptide immunogen containing luteinizing hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer.
XX
PS Claim 9; Page 71; 102pp; English.
XX
CC The present sequence represents a peptide immunogen comprising an invasin
CC domain immunostimulatory peptide of Yersinia sp., a synthetic helper T
CC cell (Th) epitope and a target antigen, luteinising hormone-releasing
CC hormone (LHRH). The synthetic Th epitope is derived from a structured
CC synthetic antigen library (SSAL) designated SSAL1 Th1. SSAL Th1 is
CC modelled after a promiscuous epitope taken from the F protein of the
CC Measles virus. The peptide immunogens cause induction of a specific
CC immune response to LHRH which is involved in regulation of
CC spermatogenesis, ovulation, oestrus, sexual development and secretion of
CC sex hormones. Provision of a promiscuous T helper epitope (which is
CC functional in genetically diverse subjects) provides optimum
CC immunogenicity to the B cell epitopes of the target antigen and thus high
CC antibody titres against the target antigen. The peptide immunogens of the
CC invention are used to vaccinate against mammalian LHRH, for use as
CC (reversible) contraceptive; control of hormone-dependent tumours (cancer
CC of prostate or breast, also endometriosis); to prevent boar taint (and
CC improve meat quality) and for immunocastration
XX
SQ Sequence 45 AA;

Query Match 67.5%; Score 108; DB 3; Length 45;
Best Local Similarity 71.4%; Pred. No. 8.2e-09;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```
Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||::||: |   ||||||||
Db         19 ISEIKGVIVHKIEGIGGE--HWSYGLRP 44
```

RESULT 15
AAY91165
ID AAY91165 standard; peptide; 45 AA.
XX
AC AAY91165;
XX

DT 12-SEP-2003 (revised)
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:45.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Measles virus.
OS Rattus sp.
OS Chimeric.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US013975.
XX
PR 20-JUN-1998; 98US-00100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus.
XX
PS Example 1; Page 80; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and immunocastration)
CC ; for promoting the growth of animals; or for treating allergies or
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in
CC genetically diverse subjects) into an immunogen improves capacity to
CC induce a strong T helper cell-mediated immune response, resulting in
CC production of antibodies against a target antigen. Th can replace carrier
CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121
CC represents a promiscuous T helper epitope from the measles virus F (MVF)
CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence
CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)
CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes
CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a
CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide
CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and
CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th
CC epitope. Somatostatin immunogens may be used to promote growth in
CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and
CC AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may
CC be used to prevent HIV infection of T cells. AAY90212 is a modified
CC version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219
CC are Th epitope/IgE CH3 antigenic peptides which may be used in the
CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth
CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this
CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum
CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.
CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are
CC immunogens comprising a CETP peptide and a Th epitope which may be used
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory
CC invasin protein epitope from Yersinia species, and hinge spacer peptide,
CC both of which may optionally be used in the antigenic peptides of the
CC invention. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 45 AA;

Query Match 67.5%; Score 108; DB 3; Length 45;
Best Local Similarity 71.4%; Pred. No. 8.2e-09;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```
Qy          3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30
              :|||||||||::||: |   ||||||||
Db         19 ISEIKGVIVHKIEGIGGE--HWSYGLRP 44
```

Search completed: March 10, 2004, 09:12:09
Job time : 47.6809 secs

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 4.71595 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-6
Perfect score: 29
Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	6	9	US-09-848-834A-6	Sequence 6, Appli
2	29	100.0	8	9	US-09-848-834A-7	Sequence 7, Appli
3	29	100.0	31	9	US-09-848-834A-15	Sequence 15, Appl
4	29	100.0	34	9	US-09-848-834A-13	Sequence 13, Appl
5	29	100.0	36	9	US-09-848-834A-16	Sequence 16, Appl
6	29	100.0	37	9	US-09-848-834A-14	Sequence 14, Appl
7	29	100.0	46	9	US-09-848-834A-19	Sequence 19, Appl
8	29	100.0	47	9	US-09-848-834A-17	Sequence 17, Appl
9	29	100.0	50	9	US-09-848-834A-18	Sequence 18, Appl
10	29	100.0	51	9	US-09-848-834A-20	Sequence 20, Appl
11	29	100.0	53	9	US-09-864-761-47649	Sequence 47649, A
12	29	100.0	75	9	US-09-731-872-425	Sequence 425, App
13	29	100.0	75	10	US-09-876-997-425	Sequence 425, App
14	29	100.0	200	9	US-09-870-756-38	Sequence 38, Appl
15	29	100.0	200	9	US-09-874-585B-38	Sequence 38, Appl
16	29	100.0	492	14	US-10-156-761-9393	Sequence 9393, Ap
17	29	100.0	1197	9	US-09-738-626-6493	Sequence 6493, Ap
18	29	100.0	3186	15	US-10-210-130-34	Sequence 34, Appl
19	29	100.0	3208	15	US-10-210-130-38	Sequence 38, Appl
20	29	100.0	3252	15	US-10-210-130-36	Sequence 36, Appl
21	29	100.0	3262	15	US-10-379-381-4	Sequence 4, Appli
22	29	100.0	3268	15	US-10-379-381-2	Sequence 2, Appli
23	27	93.1	122	10	US-09-764-891-3273	Sequence 3273, Ap
24	27	93.1	606	14	US-10-225-567A-438	Sequence 438, App
25	27	93.1	645	9	US-09-796-338A-2	Sequence 2, Appli
26	27	93.1	645	10	US-09-977-418-28	Sequence 28, Appl
27	27	93.1	645	10	US-09-977-418-46	Sequence 46, Appl
28	27	93.1	645	10	US-09-977-033A-28	Sequence 28, Appl
29	27	93.1	645	10	US-09-977-033A-46	Sequence 46, Appl
30	27	93.1	645	10	US-09-977-751C-28	Sequence 28, Appl
31	27	93.1	645	10	US-09-977-751C-46	Sequence 46, Appl
32	27	93.1	645	10	US-09-977-639A-28	Sequence 28, Appl
33	27	93.1	645	10	US-09-977-639A-46	Sequence 46, Appl
34	27	93.1	645	11	US-09-977-819B-28	Sequence 28, Appl
35	27	93.1	645	11	US-09-977-819B-46	Sequence 46, Appl
36	27	93.1	645	14	US-10-282-837-2	Sequence 2, Appli
37	27	93.1	645	14	US-10-145-586-2	Sequence 2, Appli
38	27	93.1	683	10	US-09-977-418-26	Sequence 26, Appl
39	27	93.1	683	10	US-09-977-033A-26	Sequence 26, Appl
40	27	93.1	683	10	US-09-977-751C-26	Sequence 26, Appl
41	27	93.1	683	10	US-09-977-639A-26	Sequence 26, Appl
42	27	93.1	683	11	US-09-977-819B-26	Sequence 26, Appl
43	27	93.1	690	9	US-09-828-366-7	Sequence 7, Appli
44	27	93.1	690	9	US-09-909-320-49	Sequence 49, Appl
45	27	93.1	690	9	US-09-909-088B-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-6
; Sequence 6, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-848-834A-6

Query Match 100.0%; Score 29; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SSGPSL 6
              ||||||
Db          1 SSGPSL 6
```

RESULT 2
US-09-848-834A-7
; Sequence 7, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens

```
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic peptide
US-09-848-834A-7

  Query Match           100.0%;  Score 29;  DB 9;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 7.1e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db          1 SSGPSL 6


RESULT 3
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
;   OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the Te
;   OTHER INFORMATION: tanus toxoid precursor (Tentoxylysin)
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(31)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-15

  Query Match           100.0%;  Score 29;  DB 9;  Length 31;
  Best Local Similarity  100.0%;  Pred. No. 77;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 4
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: he Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
;   OTHER INFORMATION: virus fusion protein, P
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

  Query Match           100.0%;  Score 29;  DB 9;  Length 34;
  Best Local Similarity  100.0%;  Pred. No. 85;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 5
US-09-848-834A-16
; Sequence 16, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 36
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the C
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
;   OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
```

```
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
;   OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
;   OTHER INFORMATION: (CSP) protein
US-09-848-834A-16

  Query Match             100.0%;  Score 29;  DB 9;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 6
US-09-848-834A-14
; Sequence 14, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-967 of
;   OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylysin)
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-14

  Query Match             100.0%;  Score 29;  DB 9;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 7
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetan
;   OTHER INFORMATION: toxoid precursor (Tentoxylysin) linked by a spacer to amino aci
;   OTHER INFORMATION: sequence 1-10 of GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(31)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
;   NAME/KEY: PEPTIDE
;   LOCATION: (32)..(37)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(46)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

  Query Match             100.0%;  Score 29;  DB 9;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 8
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid se
;   OTHER INFORMATION: uence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
```

```
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: Amidated-glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (35)..(38)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (39)..(47)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

  Query Match           100.0%;  Score 29;  DB 9;  Length 47;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 9
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer t
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match           100.0%;  Score 29;  DB 9;  Length 50;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 10
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 51
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmo
;   OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (51)..(51)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
;   OTHER INFORMATION: circumsporozoite (CSP) protein
;   NAME/KEY: PEPTIDE
;   LOCATION: (37)..(42)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (43)..(51)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

  Query Match           100.0%;  Score 29;  DB 9;  Length 51;
  Best Local Similarity  100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 11
US-09-864-761-47649
; Sequence 47649, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
```

```
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47649
;    LENGTH: 53
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: MAP TO Z83826.12
;    OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
;    OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.96
;    OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.57
;    OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.62
;    OTHER INFORMATION: SWISSPROT HIT: O52399, EVALUE 5.90e+00
;    OTHER INFORMATION: EST_HUMAN HIT: AA813575.1, EVALUE 3.00e-22
US-09-864-761-47649

  Query Match             100.0%;  Score 29;  DB 9;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db          5 SSGPSL 10


RESULT 12
US-09-731-872-425
; Sequence 425, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Jobert, Severin
;  TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
;  FILE REFERENCE: 78.US3.REG
;  CURRENT APPLICATION NUMBER: US/09/731,872
;  CURRENT FILING DATE:  2000-12-07
;  PRIOR APPLICATION NUMBER: US 60/169,629
;  PRIOR FILING DATE: 1999-12-08
;  PRIOR APPLICATION NUMBER: US 60/187,470
;  PRIOR FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 482
;  SOFTWARE: Patent.pm
; SEQ ID NO 425
;    LENGTH: 75
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-731-872-425

  Query Match             100.0%;  Score 29;  DB 9;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         41 SSGPSL 46


RESULT 13
US-09-876-997-425
; Sequence 425, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Jobert, Severin
;  TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
;  FILE REFERENCE: 78.US4.CIP
;  CURRENT APPLICATION NUMBER: US/09/876,997
;  CURRENT FILING DATE:  2001-06-08
;  PRIOR APPLICATION NUMBER: US 09/731,872
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: US 60/187,470
;  PRIOR FILING DATE: 2000-03-06
;  PRIOR APPLICATION NUMBER: US 60/169,629
;  PRIOR FILING DATE: 1999-12-08
;  NUMBER OF SEQ ID NOS: 482
;  SOFTWARE: Patent.pm
; SEQ ID NO 425
;    LENGTH: 75
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-876-997-425

  Query Match             100.0%;  Score 29;  DB 10;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         41 SSGPSL 46


RESULT 14
US-09-870-756-38
; Sequence 38, Application US/09870756
; Patent No. US20020052023A1
;   GENERAL INFORMATION:
;        APPLICANT: VIITANEN, PAUL VEIKKO
;                   BACOT, KAREN ONLEY
;                   JORDAN, DOUGLAS BRIAN
;        TITLE OF INVENTION: LUMAZINE SYNTHASE AND
;                            RIBOFLAVIN SYNTHASE
```

; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/870,756
; FILING DATE: 31-May-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: M. grisea LS
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-870-756-38

Query Match 100.0%; Score 29; DB 9; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 79 SSGPSL 84

RESULT 15
US-09-874-585B-38
; Sequence 38, Application US/09874585B
; Patent No. US20020127670A1
; GENERAL INFORMATION:
; APPLICANT: Viitanen, Paul Veikko
; APPLICANT: Jordan, Douglas Brain
; APPLICANT: Bacot, Karen Onley
; TITLE OF INVENTION: Riboflavin Synthase Genes and Enzymes and Methods of Use
; FILE REFERENCE: CL1083 US DIV2
; CURRENT APPLICATION NUMBER: US/09/874,585B
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/874,585
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 08/912,218
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Magnaporthe grisea
US-09-874-585B-38

Query Match 100.0%; Score 29; DB 9; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 79 SSGPSL 84

Search completed: March 10, 2004, 10:25:47
Job time : 4.71595 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 9.03502 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-6
Perfect score: 29
Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	6	5	AAU11417	Aau11417 Synthetic
2	29	100.0	8	5	AAU11418	Aau11418 Synthetic
3	29	100.0	31	5	AAU11426	Aau11426 Synthetic
4	29	100.0	34	5	AAU11424	Aau11424 Synthetic
5	29	100.0	36	5	AAU11427	Aau11427 Synthetic
6	29	100.0	37	5	AAU11425	Aau11425 Synthetic
7	29	100.0	46	5	AAU11430	Aau11430 Synthetic
8	29	100.0	47	5	AAU11428	Aau11428 Synthetic
9	29	100.0	50	5	AAU11429	Aau11429 Synthetic
10	29	100.0	51	3	AAG03447	Aag03447 Human sec
11	29	100.0	51	5	AAU11431	Aau11431 Synthetic
12	29	100.0	53	4	AAM18164	Aam18164 Peptide #
13	29	100.0	53	4	ABB37195	Abb37195 Peptide #
14	29	100.0	53	4	AAM70325	Aam70325 Human bon
15	29	100.0	53	4	AAM05786	Aam05786 Peptide #
16	29	100.0	75	4	AAG89305	Aag89305 Human sec
17	29	100.0	117	4	AAO10754	Aao10754 Human pol
18	29	100.0	128	2	AAY73871	Aay73871 Human pro
19	29	100.0	172	4	AAU42353	Aau42353 Propionib
20	29	100.0	172	6	ABM38872	Abm38872 Propionib
21	29	100.0	181	4	AAU30732	Aau30732 Novel hum
22	29	100.0	187	4	ABG19555	Abg19555 Novel hum
23	29	100.0	187	4	ABG19556	Abg19556 Novel hum
24	29	100.0	200	3	AAY97393	Aay97393 Magnaport
25	29	100.0	200	4	AAB04160	Aab04160 Magnoport
26	29	100.0	200	5	AAE20533	Aae20533 Magnaport
27	29	100.0	200	5	AAE21539	Aae21539 Magnaport
28	29	100.0	200	5	ABB79071	Abb79071 Magnaport
29	29	100.0	200	6	ABU11300	Abu11300 M. grisea
30	29	100.0	213	4	AAU27692	Aau27692 Human ful
31	29	100.0	227	4	AAU27864	Aau27864 Human con
32	29	100.0	280	4	AAB72381	Aab72381 Human hai
33	29	100.0	291	7	ADC12728	Adc12728 Human GPC
34	29	100.0	297	6	ABR41335	Abr41335 Human DIT
35	29	100.0	1035	4	AAM79726	Aam79726 Human pro
36	29	100.0	1062	4	AAM78742	Aam78742 Human pro
37	29	100.0	1197	4	AAG92739	Aag92739 C glutami
38	29	100.0	1356	6	ABJ19406	Abj19406 Human int
39	29	100.0	3186	7	ADE47672	Ade47672 Human NOV
40	29	100.0	3208	7	ADE47676	Ade47676 Human NOV
41	29	100.0	3252	7	ADE47674	Ade47674 Human NOV
42	29	100.0	3262	7	ADB79961	Adb79961 Mouse ser
43	29	100.0	3267	7	ADC99075	Adc99075 Human KPP
44	29	100.0	3268	7	ADB79959	Adb79959 Human kin
45	29	100.0	5533	4	ABB65772	Abb65772 Drosophil

ALIGNMENTS

RESULT 1
AAU11417
ID AAU11417 standard; peptide; 6 AA.
XX
AC AAU11417;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic spacer peptide #2.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
KW spacer peptide.
XX
OS Synthetic.
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 10; Page 6; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),

```
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   spacer peptide used in the immunogen of the invention
XX
SQ   Sequence 6 AA;

  Query Match           100.0%;  Score 29;  DB 5;  Length 6;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db          1 SSGPSL 6


RESULT 2
AAU11418
ID   AAU11418 standard; peptide; 8 AA.
XX
AC   AAU11418;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic spacer peptide #3.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer;
KW   spacer peptide.
XX
OS   Synthetic.
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 10; Page 6; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   spacer peptide used in the immunogen of the invention
XX
SQ   Sequence 8 AA;

  Query Match           100.0%;  Score 29;  DB 5;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db          1 SSGPSL 6


RESULT 3
AAU11426
ID   AAU11426 standard; peptide; 31 AA.
XX
AC   AAU11426;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 7.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .31
FT                   /note= "Tetanus toxoid sequence (830-844 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
```

```
SQ   Sequence 31 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 85;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 4
AAU11424
ID   AAU11424 standard; peptide; 34 AA.
XX
AC   AAU11424;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 5.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Measles virus.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .34
FT                   /note= "Measles virus fusion protein F epitope"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 34 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 5
AAU11427
ID   AAU11427 standard; peptide; 36 AA.
XX
AC   AAU11427;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 8.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .36
FT                   /note= "Malaria CSP protein (378-398 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
```

```
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 36 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 6
AAU11425
ID   AAU11425 standard; peptide; 37 AA.
XX
AC   AAU11425;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 6.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid sequence (947-967 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 37 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 7
AAU11430
ID   AAU11430 standard; peptide; 46 AA.
XX
AC   AAU11430;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 11.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .31
FT                   /note= "Tetanus toxoid (830-844 aa)"
FT   Peptide         32. .37
FT                   /note= "Spacer peptide"
FT   Peptide         38. .46
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   46
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
```

PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 11; Page 12; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 46 AA;

Query Match 100.0%; Score 29; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 11 SSGPSL 16

RESULT 8
AAU11428
ID AAU11428 standard; peptide; 47 AA.
XX
AC AAU11428;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 9.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope (1. .10
FT aa)"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11. .16
FT /note= "Spacer peptide"
FT Peptide 17. .34
FT /note= "Malaria CSP protein (288-302 aa)"
FT Peptide 35. .38
FT /note= "Spacer peptide"
FT Peptide 39. .47
FT /note= "Gonadotrophin releasing hormone epitope (2-10
FT aa)"
FT Modified-site 47
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 11; Page 11; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 47 AA;

Query Match 100.0%; Score 29; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 11 SSGPSL 16

RESULT 9
AAU11429
ID AAU11429 standard; peptide; 50 AA.
XX
AC AAU11429;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 10.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Clostridium tetani.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope (1. .10
FT aa)"
FT Misc-difference 1

```
FT                  /label= OTHER
FT                  /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide        11. .16
FT                  /note= "Spacer peptide"
FT   Peptide        17. .37
FT                  /note= "Tetanus toxoid (947-967 aa)"
FT   Peptide        38. .41
FT                  /note= "Spacer peptide"
FT   Peptide        42. .50
FT                  /note= "Gonadotrophin releasing hormone epitope (2-10
FT                  aa)"
FT   Modified-site  50
FT                  /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 50 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 10
AAG03447
ID   AAG03447 standard; protein; 51 AA.
XX
AC   AAG03447;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 7528.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC03453.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 13; SEQ ID NO 7528; 71pp + Sequence Listing; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number of
CC   5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
CC   prepared from total human RNAs or polyA+ RNAs derived from 30 different
CC   tissues. EST sequences usually correspond mainly to the 3' untranslated
CC   region (UTR) of the mRNA because they are often obtained from oligo-dT
CC   primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC   sequences derived from the 5' ends of mRNAs and even in those cases where
CC   longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC   included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC   therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC   are also used in diagnostic, forensic, gene therapy and chromosome
CC   mapping procedures. They are used to obtain upstream regulatory sequences
CC   and to design expression and secretion vectors
XX
SQ   Sequence 51 AA;

  Query Match             100.0%;  Score 29;  DB 3;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         16 SSGPSL 21


RESULT 11
AAU11431
ID   AAU11431 standard; peptide; 51 AA.
XX
AC   AAU11431;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 12.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
```

```
FT                  /label= OTHER
FT                  /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide        11. .16
FT                  /note= "Spacer peptide"
FT   Peptide        17. .36
FT                  /note= "Malaria CSP protein (378-398 aa)"
FT   Peptide        37. .42
FT                  /note= "Spacer peptide"
FT   Peptide        43. .51
FT                  /note= "Gonadotrophin releasing hormone epitope (2-10
FT                  aa)"
FT   Modified-site  51
FT                  /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 12-13; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 51 AA;

  Query Match             100.0%;  Score 29;  DB 5;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         11 SSGPSL 16


RESULT 12
AAM18164
ID   AAM18164 standard; protein; 53 AA.
XX
AC   AAM18164;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Peptide #4598 encoded by probe for measuring cervical gene expression.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 27; SEQ ID NO 22990; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC   by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC   can be used to produce a single exon microarray, which can be used for
CC   measuring human gene expression in a sample derived from human cervical
CC   epithelial cells. By measuring gene expression, the probes are therefore
CC   useful in grading and/or staging of diseases of the cervix, notably
CC   cervical cancer. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 53 AA;

  Query Match             100.0%;  Score 29;  DB 4;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db          5 SSGPSL 10


RESULT 13
ABB37195
ID   ABB37195 standard; peptide; 53 AA.
XX
AC   ABB37195;
XX
DT   04-FEB-2002  (first entry)
XX
DE   Peptide #4701 encoded by human foetal liver single exon probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
```

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 29830; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 29; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 5 SSGPSL 10

RESULT 14
AAM70325
ID AAM70325 standard; protein; 53 AA.
XX
AC AAM70325;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30631.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 30631; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 29; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 5 SSGPSL 10

RESULT 15
AAM05786
ID AAM05786 standard; protein; 53 AA.
XX
AC AAM05786;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #4468 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 27; SEQ ID NO 14526; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 53 AA;

Query Match 100.0%; Score 29; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SSGPSL 6
              ||||||
Db          5 SSGPSL 10
```

Search completed: March 10, 2004, 09:12:08
Job time : 11.035 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 5.92996 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-6
Perfect score: 29
Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	113	16	Q83F79	Q83f79 coxiella bu
2	29	100.0	174	5	Q8MZ01	Q8mz01 drosophila
3	29	100.0	200	3	Q9UVT8	Q9uvt8 magnaporthe
4	29	100.0	213	4	Q9UJG1	Q9ujg1 homo sapien
5	29	100.0	213	11	Q8VEL0	Q8vel0 mus musculu
6	29	100.0	214	11	Q9D8Y9	Q9d8y9 mus musculu
7	29	100.0	215	13	Q7SYQ6	Q7syq6 xenopus lae
8	29	100.0	226	10	Q9FJX4	Q9fjx4 arabidopsis
9	29	100.0	257	16	Q8VKE5	Q8vke5 mycobacteri
10	29	100.0	277	4	Q8IXV0	Q8ixv0 homo sapien
11	29	100.0	284	2	Q9L913	Q9l913 aeromonas v
12	29	100.0	348	16	Q8YJL9	Q8yjl9 brucella me
13	29	100.0	358	16	Q7U3U5	Q7u3u5 synechococc
14	29	100.0	392	10	Q7XK21	Q7xk21 oryza sativ
15	29	100.0	395	4	Q8N4F2	Q8n4f2 homo sapien
16	29	100.0	395	12	Q86606	Q86606 simian viru
17	29	100.0	405	16	Q97H14	Q97h14 clostridium
18	29	100.0	430	5	Q7YVE4	Q7yve4 trypanosoma
19	29	100.0	437	16	Q8XKM3	Q8xkm3 clostridium
20	29	100.0	486	13	Q7T0R3	Q7t0r3 xenopus lae
21	29	100.0	492	16	Q82M06	Q82m06 streptomyce
22	29	100.0	507	16	Q8FY77	Q8fy77 brucella su
23	29	100.0	520	13	Q7ZUA3	Q7zua3 brachydanio
24	29	100.0	532	16	P72605	P72605 synechocyst
25	29	100.0	593	16	Q8A6C8	Q8a6c8 bacteroides
26	29	100.0	615	11	Q8BP38	Q8bp38 mus musculu
27	29	100.0	1009	16	Q8PM63	Q8pm63 xanthomonas
28	29	100.0	1074	10	Q9SYZ6	Q9syz6 arabidopsis
29	29	100.0	1197	16	Q8NM57	Q8nm57 corynebacte
30	29	100.0	2078	16	Q98K31	Q98k31 rhizobium l
31	29	100.0	3262	11	Q9EQJ5	Q9eqj5 mus musculu
32	27	93.1	103	16	Q89BU2	Q89bu2 bradyrhizob
33	27	93.1	121	10	Q9SY80	Q9sy80 arabidopsis
34	27	93.1	123	11	Q9CVW5	Q9cvw5 mus musculu
35	27	93.1	183	5	Q8SW89	Q8sw89 encephalito
36	27	93.1	189	11	Q99PH9	Q99ph9 mus musculu
37	27	93.1	189	11	Q99PI3	Q99pi3 mus musculu
38	27	93.1	192	10	Q9SY78	Q9sy78 arabidopsis
39	27	93.1	207	10	Q9FUS7	Q9fus7 arabidopsis
40	27	93.1	218	10	Q9SXZ1	Q9sxz1 daucus caro
41	27	93.1	222	5	O01786	O01786 caenorhabdi
42	27	93.1	227	10	Q9FUS8	Q9fus8 arabidopsis
43	27	93.1	227	10	Q94II0	Q94ii0 arabidopsis
44	27	93.1	227	10	Q9FUS9	Q9fus9 arabidopsis
45	27	93.1	227	11	Q8VCG7	Q8vcg7 mus musculu

ALIGNMENTS

RESULT 1
Q83F79
ID Q83F79 PRELIMINARY; PRT; 113 AA.
AC Q83F79;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN CBU0069.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
RT burnetii.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL; AE016960; AAO89636.1; -.
DR TIGR; CBU0069; -.
DR InterPro; IPR002110; ANK.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 113 AA; 12752 MW; 5C82A677DEE7BD95 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 113;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
||||||
Db 35 SSGPSL 40

```
RESULT 2
Q8MZ01
ID   Q8MZ01      PRELIMINARY;      PRT;   174 AA.
AC   Q8MZ01;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RE33542p.
GN   BCDNA:RE33542.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY113451; AAM29456.1; -.
DR   FlyBase; FBgn0063034; BcDNA:RE33542.
SQ   SEQUENCE   174 AA;  19281 MW;  05C5CE1AB43DB8DA CRC64;

  Query Match             100.0%;  Score 29;  DB 5;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db        122 SSGPSL 127


RESULT 3
Q9UVT8
ID   Q9UVT8      PRELIMINARY;      PRT;   200 AA.
AC   Q9UVT8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   6,7-dimethyl-8-ribityllumazine synthase.
OS   Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX   NCBI_TaxID=148305;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Persson K., Schneider G., Jordan D.B., Viitanen P.V., Sandalova T.;
RT   "Comparison of the crystal structures of the pentameric fungal and the
RT   icosahedral plant lumazine synthases.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF148449; AAD55372.1; -.
DR   PDB; 1C41; 07-AUG-00.
DR   GO; GO:0009349; C:riboflavin synthase complex; IEA.
DR   GO; GO:0004746; F:riboflavin synthase activity; IEA.
DR   GO; GO:0009231; P:vitamin B2 biosynthesis; IEA.
DR   InterPro; IPR002180; DMRL_synthase.
DR   Pfam; PF00885; DMRL_synthase; 1.
DR   ProDom; PD003664; DMRL_synthase; 1.
DR   TIGRFAMs; TIGR00114; ribH; 1.
SQ   SEQUENCE   200 AA;  21072 MW;  E079F049D32271B5 CRC64;

  Query Match             100.0%;  Score 29;  DB 3;  Length 200;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db         79 SSGPSL 84


RESULT 4
Q9UJG1
ID   Q9UJG1      PRELIMINARY;      PRT;   213 AA.
AC   Q9UJG1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein dJ473b4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rhodes S., Huckle E.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL137163; CAB69662.1; -.
DR   EMBL; BC005700; AAH05700.1; -.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000535; MSP_domain.
DR   InterPro; IPR008962; PapD-like.
DR   Pfam; PF00635; MSP_domain; 1.
DR   PROSITE; PS50202; MSP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   213 AA;  24086 MW;  2E0944F48058B47F CRC64;

  Query Match             100.0%;  Score 29;  DB 4;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SSGPSL 6
              ||||||
Db        155 SSGPSL 160


RESULT 5
Q8VEL0
ID   Q8VEL0      PRELIMINARY;      PRT;   213 AA.
AC   Q8VEL0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RIKEN cDNA 1810018L05 gene.
GN   1810018L05RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC018329; AAH18329.1; -.
DR   MGD; MGI:1917630; 1810018L05Rik.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000535; MSP_domain.
DR   InterPro; IPR008962; PapD-like.
DR   Pfam; PF00635; MSP_domain; 1.
DR   PROSITE; PS50202; MSP; 1.
SQ   SEQUENCE   213 AA;  24074 MW;  FB72756A0528CB34 CRC64;

  Query Match             100.0%;  Score 29;  DB 11;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```